

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 02:35:54 ; Search time 3265 Seconds
(without alignments)
12136.941 Million cell updates/sec

Title: US-09-927-091-3_COPY_2500_3826
Perfect score: 1327
Sequence: I tccacagtgtcacaggtag.....aaaaaaaaaaaaaaaa 1327

Scoring table: OHGAC NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database : EST.*

- 1: em_estba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estnu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_htc.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_htc.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: em_gss_hum.*
- 18: em_gss_inv.*
- 19: em_gss_pin.*
- 20: em_gss_vrt.*
- 21: em_gss_fun.*
- 22: em_gss_mam.*
- 23: em_gss_mus.*
- 24: em_gss_pro.*
- 25: em_gss_rod.*
- 26: em_gss_pbg.*
- 27: em_gss_vrl.*
- 28: gb_gss1.*
- 29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	751	56.6	835	14	CA310925
C 2	721	54.3	769	14	CA444761
C 3	688	51.8	1049	12	BM457033
C 4	656	49.4	781	14	CB956370

C 5	634	47.8	634	12	BQ018441
C 6	632	47.6	632	12	BM994555
C 7	587	44.2	682	14	CA309985
C 8	582	43.9	633	12	BM994326
C 9	519	39.1	519	9	AL138362
C 10	496	37.4	510	13	EX283437
C 11	489	36.9	501	9	AL138363
C 12	455	34.3	470	10	AW515028
C 13	440	33.2	486	9	AI457621
C 14	434	32.7	451	14	CA434008
C 15	431	32.5	477	10	BE671191
C 16	423	31.9	557	9	AUL54016
C 17	421	31.7	423	9	AI066431
C 18	421	31.7	501	13	BQ187235
C 19	415	31.3	983	10	BE795637
C 20	405	30.5	456	10	BE856182
C 21	403	30.4	456	9	AA700789
C 22	399	30.1	502	14	CK300567
C 23	391	29.5	512	9	AI681374
C 24	378	28.5	431	14	R71654
C 25	370	27.9	497	10	BE244684
C 26	319	24.0	424	14	CA434086
C 27	305	23.0	543	14	R71157
C 28	304	22.9	304	9	AI383065
C 29	299	22.5	376	9	BE245533
C 30	286	21.6	385	9	AA644653
C 31	285	21.5	434	10	AW139822
C 32	283	21.3	431	14	H47130
C 33	280	21.3	440	14	H78426
C 34	280	21.1	513	9	AA037876
C 35	277	20.9	543	9	AW070327
C 36	269	20.3	389	14	T85444
C 37	269	20.3	470	10	BE219390
C 38	260	19.6	337	10	BE245721
C 39	255	19.2	749	10	BF696473
C 40	247	18.6	270	13	BU754096
C 41	241	18.2	378	10	AW964518
C 42	226	17.0	350	9	AA191674
C 43	225	17.0	334	10	BE817016
C 44	221	16.7	405	12	BG912772
C 45	221	16.7	768	9	AL554421
C 46	212	16.0	931	13	BU527114
C 47	211	15.9	310	9	AI864353
C 48	196	14.8	270	12	BQ025918
C 49	194	14.6	488	9	AA037877
C 50	182	13.7	401	12	BM678923
C 51	181	13.6	295	13	BQ367558
C 52	180	13.6	346	9	AA909074
C 53	164	12.4	342	9	AI242830
C 54	163	12.3	295	14	H78226
C 55	153	11.5	363	9	AA320033
C 56	150	11.3	1129	12	BM547517
C 57	138	10.4	153	12	BQ022505
C 58	124	9.3	269	10	BQ223505
C 59	113	8.5	304	10	BQ841308
C 60	110	8.3	593	9	BE246494
C 61	105	7.9	135	12	BQ022575
C 62	100	7.5	901	13	BX348422
C 63	84	6.3	334	12	BQ025276
C 64	74	5.6	186	14	Z41723
C 65	70	5.3	906	13	BU171407
C 66	67	5.0	341	10	BF987119
C 67	62	4.7	480	14	T85347
C 68	43	3.2	732	9	AL554364
C 69	41	3.1	432	14	CF117086
C 70	41	3.1	551	10	BE758068
C 71	41	3.1	552	10	BE753612
C 72	35	2.6	300	14	Z44845
C 73	34	2.6	394	12	BI899337
C 74	31	2.3	550	10	BE753490
C 75	30	2.3	312	14	CF116526
C 76	30	2.3	524	14	CF116527
C 77	29	2.2	167	14	CF313132

C 78	28	2.1	154	13	BQ792699	BQ792699 EST 8419
C 79	28	2.1	171	12	BG688740	BG688740 336481 BA
C 80	28	2.1	199	10	BF463886	BF463886 UI-M-CG0P
C 81	28	2.1	234	13	BU954379	BU954379 AGENCOURT
C 82	28	2.1	342	12	BI300710	BI300710 UI-R-CV2-
C 83	28	2.1	351	9	AA184228	AA184228 mc32g04.f
C 84	28	2.1	358	12	BI300041	BI300041 UI-R-CV2-
C 85	28	2.1	406	12	BI297552	BI297552 UI-R-CV2-
C 86	28	2.1	464	12	BI298701	BI298701 UI-R-CV2-
C 87	28	2.1	470	10	BF776486	BF776486 287238 MA
C 88	28	2.1	478	10	BF457922	BF457922 UI-M-BZ1-
C 89	28	2.1	488	12	BG373733	BG373733 UI-R-CV1-
C 90	28	2.1	518	10	BE996205	BE996205 UI-M-BZ1-
C 91	28	2.1	519	9	AI227813	AI227813 EST224508
C 92	28	2.1	560	12	BM900146	BM900146 UI-M-DJ1-
C 93	28	2.1	623	12	BI300357	BI300357 UI-R-CV2-
C 94	28	2.1	652	28	BH684258	BH684258 BOMM30TF
C 95	28	2.1	774	29	CC536668	CC536668 CH240.415
C 96	28	2.1	1086	14	CK231937	CK231937 ILLUMIGEN
C 97	27	2.0	72	10	AW169355	AW169355 x122g08.x
C 98	27	2.0	117	14	CF201671	CF201671 RR890915N
C 99	27	2.0	154	12	BI003499	BI003499 MK3-HN012
C 100	27	2.0	209	14	CB821576	CB821576 EST 2427
C 101	27	2.0	210	12	BI845801	BI845801 fs98b02.x
C 102	27	2.0	229	12	BI845510	BI845510 fs98b02.y
C 103	27	2.0	346	10	BF191229	BF191229 237997 MA
C 104	27	2.0	349	10	BE674222	BE674222 d77a06.x
C 105	27	2.0	353	13	BQ172169	BQ172169 WHE2003.A
C 106	27	2.0	358	10	BF191683	BF191683 239337 MA
C 107	27	2.0	369	9	AA280737	AA280737 zs96h10.s
C 108	27	2.0	372	10	AW138882	AW138882 UI-H-B1L-
C 109	27	2.0	396	12	BM027121	BM027121 GIT000041
C 110	27	2.0	405	14	CA336437	CA336437 NISC_lv01
C 111	27	2.0	440	29	CE533755	CE533755 t1gr-gss-
C 112	27	2.0	448	12	BM874938	BM874938 if39d08.x
C 113	27	2.0	477	10	BF729120	BF729120 100007F0
C 114	27	2.0	487	14	CA663515	CA663515 wlmk1.pk0
C 115	27	2.0	493	10	AW129612	AW129612 xel19g11.x
C 116	27	2.0	501	14	CB129334	CB129334 K-BST0178
C 117	27	2.0	566	10	AW118907	AW118907 xd9eb11.x
C 118	27	2.0	587	9	AI676005	AI676005 wc04b01.x
C 119	27	2.0	587	14	CA814536	CA814536 CA48LNI01
C 120	27	2.0	607	14	CB343776	CB343776 CA32EN000
C 121	27	2.0	636	29	CE637121	CE637121 t1gr-gss-
C 122	27	2.0	710	10	BF540944	BF540944 603067582
C 123	27	2.0	711	14	CA759041	CA759041 BR060014B
C 124	27	2.0	713	14	CD373920	CD373920 WHE2613.E
C 125	27	2.0	714	9	AU068118	AU068118 AU068118
C 126	27	2.0	768	14	CB343742	CB343742 CA32EN000
C 127	27	2.0	1026	13	BX347228	BX347228 BX347228
C 128	27	2.0	1044	14	CD502799	CD502799 CDA58-P12
C 129	26	2.0	54	14	CF327511	CF327511 NACH--02-
C 130	26	2.0	96	14	CF327512	CF327512 NACH--02-
C 131	26	2.0	101	13	BQ794536	BQ794536 EST 3474
C 132	26	2.0	115	14	CF314604	CF314604 HD--03-D0
C 133	26	2.0	118	14	CF204241	CF204241 RR890915I
C 134	26	2.0	164	10	BE946330	BE946330 UI-M-BZ0-
C 135	26	2.0	189	12	BI941630	BI941630 sc80f06.y
C 136	26	2.0	191	14	CF328196	CF328196 NACL--02-
C 137	26	2.0	217	12	BI798001	BI798001 H094H04.E
C 138	26	2.0	224	14	CB316674	CB316674 AGENCOURT
C 139	26	2.0	233	12	BI503458	BI503458 BBL70031B
C 140	26	2.0	245	12	CF575205	CF575205 MCSA1440C
C 141	26	2.0	275	12	BI134712	BI134712 UI-M-BH3-
C 142	26	2.0	279	9	AW021059	AW021059 df18c08.y
C 143	26	2.0	297	9	AA598544	AA598544 ae39b01.s
C 144	26	2.0	310	9	AA598544	AA598544 ae39b01.s
C 145	26	2.0	321	10	AW323204	AW323204 fj12c10.x
C 146	26	2.0	321	14	H57960	H57960 yr05a02.sl
C 147	26	2.0	329	14	N75931	N75931 za42a01.sl
C 148	26	2.0	348	12	BG377618	BG377618 UI-R-CU0-
C 149	26	2.0	351	10	BE123282	BE123282 945040D09
C 150	26	2.0	364	10	BF729569	BF729569 1000079F0

ALIGNMENTS

RESULT 1
CA310925 835 bp mRNA linear EST 04-NOV-2002
LOCUS UI-CF-FNO-afb-j-06-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone
DEFINITION UI-CF-FNO-afb-j-06-0-UI 3', mRNA sequence.

ACCESSION CA310925
VERSION 1
KEYWORDS CA310925.1 GI.24529023
SOURCE EST.
ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 835)
AUTHORS Bernaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT 8889548
Contact: McCray, PB
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171

FEATURES
Location/Qualifiers
1..835
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-FNO-afb-j-06-0-UI"
/tissue_type="Human Lung Epithelial cells"
/lab_host="VH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-FNO"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-FNO is a subtracted cDNA library derived from two normalized Human lung epithelial cell libraries (EN1 and D01) The library was subtracted according to according to Bernaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact: bento-soares@uiowa.edu
TAG TISSUE=Human Lung Epithelial Cell lines untreated LPS
6hr to LPS 24h
TAG LIB=UI-CF-FNO
TAG_SEQ=CTGCTCAGGT"

ORIGIN

Query Match 56.6%; Score 751; DB 14; Length 835;
Best Local Similarity 100.0%; Pred. No. 2.8e-123;
Matches 751; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 575 ATTTCCAGGAATCCAGGCCAGATTCATGTCGTGTGTGTGACAGGAAGTCCCC 634
Db 751 ATTTCCAGGAATCCAGGCCAGATTCATGTCGTGTGTGTGACAGGAAGTCCCC 692
Qy 635 CATCTGAGGAAGCACTATGCCAGAAAGTGTGACTGCAGAACTAGGCTCCCTCTGCC 694


```

QY 1183 ATTGTGAGCCTACCCCAACCTTCACTACCAAGATCTGGCCACCCAGCAGTATTTT 1242
Db 136 ATTGTGAGCCTACCCCAACCTTCACTACCAAGATCTGGCCACCCAGCAGTATTTT 77
QY 1243 TATTAAAGTTGGCCATTTATGAGTTATGATCAATTTGTATTAATTAAGTTACAG 1302
Db 76 TATTAAAGTTGGCCATTTATGAGTTATGATCAATTTGTATTAATTAAGTTACAG 17
QY 1303 A 1303
Db 16 A 16

RESULT 3
BM457033
LOCUS BM457033 1049 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6406661 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5583269
5', mRNA sequence.
ACCESSION BM457033.1 GI:18506073
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1049)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12346 row: b column: 06
High quality sequence stop: 671.
FEATURES
source
1..1049
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5583269"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dr primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Query Match 51.8%; Score 688; DB 12; Length 1049;
Best Local Similarity 100.0%; Pred. No. 2.4e-112;
Matches 688; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCACAGTGGTCACAGTAGTACCTGGTCCCTAGGGTGCCTGAGAGCAACCTCTCCCTGC 60
Db 115 TCCACAGTGGTCACAGTAGTACCTGGTCCCTAGGGTGCCTGAGAGCAACCTCTCCCTGC 174
QY 61 CACCCCCCACCAAGAACTATATGTTCTCTACTTCTCCCACTGATCTCTGTCAGTGTAT 120
Db 175 CACCCCCCACCAAGAACTATATGTTCTCTACTTCTCCCACTGATCTCTGTCAGTGTAT 234
QY 121 GATGTCGTGGCTGTGGAGGACCTGGTAGTTAGTTCACACATATAGTATGTGCA 180
Db 235 GATGTCGTGGCTGTGGAGGACCTGGTAGTTAGTTCACACATATAGTATGTGCA 294

```

```

QY 181 CCACCTTCTCTGCCACAGCCGAGGACAGGGTGAAGGTATACCCAAAGCTGATGCGAG 240
Db 295 CCACCTTCTCTGCCACAGCCGAGGACAGGGTGAAGGTATACCCAAAGCTGATGCGAG 354
QY 241 CCATTAGCTTAAAGCACTGAGGACAAGCCCTCCCTGATGATCGAGGTCCCACTAG 300
Db 355 CCATTAGCTTAAAGCACTGAGGACAAGCCCTCCCTGATGATCGAGGTCCCACTAG 414
QY 301 CTCTGACAGAGTCTCAGCCCAACCTCTTTCAGCCAGGSCCTCTGTGACCTGCTAGGGTGCA 360
Db 415 CTCTGACAGAGTCTCAGCCCAACCTCTTTCAGCCAGGSCCTCTGTGACCTGCTAGGGTGCA 474
QY 361 GGAGGCTTCCAGAGCAGTTGTTGTAATTAGGACCCAAAGCACTGGGAGGGGCTGTTGGCT 420
Db 475 GGAGGCTTCCAGAGCAGTTGTTGTAATTAGGACCCAAAGCACTGGGAGGGGCTGTTGGCT 534
QY 421 AGACCCCTTCTGAGACTTGGCANTCTATCTCAGTTAGGATCCTGTCGAGAAACAAAGAGC 480
Db 535 AGACCCCTTCTGAGACTTGGCANTCTATCTCAGTTAGGATCCTGTCGAGAAACAAAGAGC 594
QY 481 CACTGTAGCTGGTTTAAATTAGACAAGGATTTACTACTGGCCCTCTGTGGCTTGCAAAA 540
Db 595 CACTGTAGCTGGTTTAAATTAGACAAGGATTTACTACTGGCCCTCTGTGGCTTGCAAAA 654
QY 541 TTGTGGAAGAGCTGGAGAGCAGTCTCTGATTTCCAGGAATCCAGAGCCCGAGAT 600
Db 655 TTGTGGAAGAGCTGGAGAGCAGTCTCTGATTTCCAGGAATCCAGAGCCCGAGAT 714
QY 601 TCATCATGCTGTTGTGACAGGAAAGCTGCCCCCATCTGCAGAGCCCACTATGCCAGA 660
Db 715 TCATCATGCTGTTGTGACAGGAAAGCTGCCCCCATCTGCAGAGCCCACTATGCCAGA 774
QY 661 AAGCTGCTGACTGCAGAACTAGGCTCCC 688
Db 775 AAGCTGCTGACTGCAGAACTAGGCTCCC 802

RESULT 4
CB956370 781 bp mRNA linear EST 29-APR-2003
LOCUS AGENCOURT_13666719 NIH_MGC_184 Homo sapiens cDNA clone
DEFINITION IMAGE:30353395 5', mRNA sequence.
ACCESSION CB956370.1 GI:30212487
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 781)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM154 row: a column: 20
High quality sequence stop: 582.
FEATURES
source
1..781
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30353395"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_184"
/note="Organ: Pooled-glandular; Vector: pDNR-LIB; Site_1:

```


QY 1052 GTCTGGGGCTAGATTCTGCACCTTGGGGTCTCTGACACACACACCATCCCAAGTAGC 1111
 Db 274 GTCTGGGGCTAGATTCTGCACCTTGGGGTCTCTGACACACACACCATCCCAAGTAGC 215
 QY 1112 CGGAAGAGCTAAACACAGGGGGTCTTAAATGGCTGCCCGCCACCCGGGCTCCCTT 1171
 Db 214 CGGAAGAGCTAAACACAGGGGGTCTTAAATGGCTGCCCGCCACCCGGGCTCCCTT 155
 QY 1172 GGGCAAAAGAAATGTGAGCCCTACCCCAACCCCTTCAACTACCAAGATCTGGGCCACCCC 1231
 Db 154 GGGCAAAAGAAATGTGAGCCCTACCCCAACCCCTTCAACTACCAAGATCTGGGCCACCCC 95
 QY 1232 AGCAGTATTTTATTTAAATGTTGCCCATTTTATGATGATGATGATGATGATGATTAAT 1291
 Db 94 AGCAGTATTTTATTTAAATGTTGCCCATTTTATGATGATGATGATGATGATTAAT 35
 QY 1292 TAAAGTTACAGATGTCAAAAAAATAAAAAA 1325
 Db 34 TAAAGTTACAGATGTCAAAAAAATAAAAAA 1

RESULT 6
 BM994555/c
 LOCUS
 DEFINITION UI-H-DHO-aum-m-09-0-UI-s1 NCI CGAP_DHO Homo sapiens cDNA clone
 IMAGE:5871536 3', mRNA sequence.
 ACCESSION BM994555
 VERSION BM994555.1 GI:19719456
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 632)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Jose Mercuende
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES
 Location/Qualifiers
 1..632
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5871536"
 /tissue_type="Metastatic Chondrosarcoma"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP DHO"
 /note="Organ: Lung; Vector: p7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
 NCI CGAP DHO is a cDNA library containing the following tissue(s): VS-8 Cell line from Metastatic Chondrosarcoma in Lung. The library was constructed according to Snald, Lemmon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dr primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tag for this library is AGATCATTCG."

TAG TISSUE=lung
 TAG_LIB=UI-H-DHO
 TAG_SEQ=AGATCATTCG"

ORIGIN

Query Match 47.6%; Score 632; DB 12; Length 632;
 Best Local Similarity 100.0%; Pred. No. 2.6e-102;
 Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 694 CAGGTCGCGTCAGCAATAGATGTCCTGAGCGCTGCCCTCTCCACATTCACATGATT 753
 Db 632 CAGGTCGCGTCAGCAATAGATGTCCTGAGCGCTGCCCTCTCCACATTCACATGATT 573
 QY 754 CCCAAATCTAAATTTTACAAGAGATTCTGTTGGGGAACTTTAAGTCAGATCCAGAACC 813
 Db 572 CCCAAATCTAAATTTTACAAGAGATTCTGTTGGGGAACTTTAAGTCAGATCCAGAACC 513
 QY 814 TTGCGTCGAAGGAGTCTGGGAAATGTCATTTCCCTAGAAGAAAGTTAGGGTGGGTGGAG 873
 Db 512 TTGCGTCGAAGGAGTCTGGGAAATGTCATTTCCCTAGAAGAAAGTTAGGGTGGGTGGAG 453
 QY 874 CAAGCCCCACCTCGTTTTTCTGCCACAGCATCCAAATCGTGAAGAACTCGGAGAGGGTG 933
 Db 452 CAAGCCCCACCTCGTTTTTCTGCCACAGCATCCAAATCGTGAAGAACTCGGAGAGGGTG 393
 QY 934 GAGTCCACATCTAGGGTGTCTGCCCCCTTGGCTCTATCCCTGCCAGAGGTGGGAACTG 993
 Db 392 GAGTCCACATCTAGGGTGTCTGCCCCCTTGGCTCTATCCCTGCCAGAGGTGGGAACTG 333
 QY 994 GAGGAGTGGGTGCAAGACTGAGCCTAAATGTCTCCCCGGGCTTGACATTTCTTTCTAGT 1053
 Db 332 GAGGAGTGGGTGCAAGACTGAGCCTAAATGTCTCCCCGGGCTTGACATTTCTTTCTAGT 273
 QY 1054 CCTGGGGCTAGATTCTGCACCTTGGGTCTCTGACACACACACCATCCCAAAGTAGCGG 1113
 Db 272 CCTGGGGCTAGATTCTGCACCTTGGGTCTCTGACACACACACCATCCCAAAGTAGCGG 213
 QY 1114 GAAGAGCTAAACACAGGGGGTTCCTTAAATGGCTGCCCGCCACCCGGGCTCCCTTGG 1173
 Db 212 GAAGAGCTAAACACAGGGGGTTCCTTAAATGGCTGCCCGCCACCCGGGCTCCCTTGG 153
 QY 1174 GCAAAAGAAATGTGAGCCCTACCCCAACCCCTTCACTACAGAACTCTGGGCCACCCAG 1233
 Db 152 GCAAAAGAAATGTGAGCCCTACCCCAACCCCTTCACTACAGAACTCTGGGCCACCCAG 93
 QY 1234 CAGTATTTTATTTAAATGTTGCCCATTTTATGATGATGATGATGATGATTAATTA 1293
 Db 92 CAGTATTTTATTTAAATGTTGCCCATTTTATGATGATGATGATGATGATTAATTA 33

1294 AAGTTACAGATGTCAAAAAAATAAAAAA 1325

32 AAGTTACAGATGTCAAAAAAATAAAAAA 1

RESULT 7
 CA309985/c
 LOCUS

DEFINITION UI-H-FT1-bid-j-07-0-UI-s1 NCI CGAP FT1 Homo sapiens cDNA clone
 UI-H-FT1-bid-j-07-0-UI-3', mRNA sequence.
 CA309985
 VERSION CA309985.1 GI:24473039
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 682)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov

CA309985 682 bp mRNA linear EST 01-NOV-2002

UI-H-FT1-bid-j-07-0-UI-s1 NCI CGAP FT1 Homo sapiens cDNA clone

UI-H-FT1-bid-j-07-0-UI-3', mRNA sequence.

CA309985.1 GI:24473039

EST.

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 682)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Gary W. Hunninghake, U of I
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES source

Location/Qualifiers
1. .682
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-F11-bid-j-07-0-UI"
/tissue type="Aveolar Macrophage"
/dev stage="Adult"
/lab host="DH10B (Life Technologies)"
/clone lib="NCI CGAP_F11"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP_F11 is a normalized cDNA library constructed from
a pool of 81 RNA samples from Alveolar Macrophages
challenged with different treatments. The library was
normalized according to Bonaldo, Lennon and Soares, Genome
Research, 6:791-806, 1996. First strand cDNA synthesis was
primed with an oligo-dT primer containing a Not I site.
Double stranded cDNA was ligated to an EcoR I adaptor,
digested with Not I, and cloned directionally into
pT7T3-Pac vector. The oligonucleotide used to prime the
synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dr)18 tail. The sequence tag for this library is
GGCCATGCGG. The tissue was provided by Dr. Gary W.
Hunninghake of the University of Iowa.
TAG TISSUE=Human Lung Aveolar Macrophage
TAG LIB=UI-H-F11
TAG_SEQ=GGCCATGCGG"

ORIGIN

Query Match 44.2%; Score 587; DB 14; Length 682;
Best Local Similarity 100.0%; Pred. No. 1.8e-94;
Matches 587; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

722 TGAGGCTGCCCCCTCCCACTTCACTCAGTCCCAATCTAAATTTTACAGAGATTC 781
607 TGAGGCTGCCCCCTCCCACTTCACTCAGTCCCAATCTAAATTTTACAGAGATTC 548
782 TGTTTGGGGAAGTAACTCAGATCCAGAACCTTGGCTGCAAGGAGTCTGGGAATGTC 841
547 TGTTTGGGGAAGTAACTCAGATCCAGAACCTTGGCTGCAAGGAGTCTGGGAATGTC 488
842 ATTTCCTAGAGGAAGTAACTCAGATCCAGAACCTTGGCTGCAAGGAGTCTGGGAATGTC 901
487 ATTTCCTAGAGGAAGTAACTCAGATCCAGAACCTTGGCTGCAAGGAGTCTGGGAATGTC 428
902 GCATCAATCGTGAAGAACTCGGAGAGGTTGAGTCCATCTAGGTTGCTCCGCC 961
427 GCATCAATCGTGAAGAACTCGGAGAGGTTGAGTCCATCTAGGTTGCTCCGCC 368
962 TTGGCTCTATCCCTGCCAGAGTGGGAATCGAGAGTGGGTGCGAAGACTGAGCTAA 1021
367 TTGGCTCTATCCCTGCCAGAGTGGGAATCGAGAGTGGGTGCGAAGACTGAGCTAA 308
1022 ATGTCTCCCGCGCTGACTTTTCTTTCTAGTCCCTGGGCTAGATCTGCACTGGGGT 1081
307 ATGTCTCCCGCGCTGACTTTTCTTTCTAGTCCCTGGGCTAGATCTGCACTGGGGT 248
1082 CTCTGACACACACACACATCCCAAGTAGCCGGAAGAGCTAAACACAGGGGTTCTTAA 1141
247 CTCTGACACACACACACATCCCAAGTAGCCGGAAGAGCTAAACACAGGGGTTCTTAA 188
1142 ATGGTGTGCCCCCGCACCGCGGCTTCCCTTGGGCAAAAGGAATGTCAGCCCTACCCCAA 1201

Db 187 ATGGCTGCCCGCCGCCACCCGGGCTCCCTTGGGCAAAAGGAATGTCAGCCCTACCCCAA 128
QY 1202 CCCTTCAACTACCAAGATCTGGGCGACCCAGCAGTATTTTATTTAAATGTTCCCAT 1261
Db 127 CCCTTCAACTACCAAGATCTGGGCGACCCAGCAGTATTTTATTTAAATGTTCCCAT 68
QY 1262 TTTATGAGTTATGATCAATTTGTATTAATTAAGTTACAGATGTC 1308
Db 67 TTTATGAGTTATGATCAATTTGTATTAATTAAGTTACAGATGTC 21

RESULT 8
BM994326/c
LOCUS
DEFINITION
UI-H-DHO-aul-k-14-0-UI.s1 NCI CGAP_DHO Homo sapiens cDNA clone
IMAGE:5871109 3', mRNA sequence.
BM994326
ACCESSION
BM994326.1 GI:19719227
VERSION
EST.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 633)
REFERENCE
NCI CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
TITLE
Unpublished (1997)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES source

Location/Qualifiers
1. .633
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5871109"
/tissue type="Metastatic Chondrosarcoma"
/dev stage="Adult"
/lab host="DH10B (Life Technologies)"
/clone lib="NCI CGAP_DHO"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP_DHO is a cDNA library containing the following
tissue(s): VS-8 Cell line from Metastatic Chondrosarcoma
in Lung. The library was constructed according to Bonaldo,
Lennon and Soares, Genome Research, 6:791-806, 1996. First
strand cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dr)18 tail. The sequence tag for this
library is AGATCATTCG.
TAG TISSUE=lung
TAG LIB=UI-H-DHO
TAG_SEQ=AGATCATTCG"

ORIGIN

Query Match 43.9%; Score 582; DB 12; Length 633;
Best Local Similarity 99.8%; Pred. No. 1.5e-93;
Matches 632; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 693 CCAGGTCCTGCCAGCAATAGATGCTCCTGAGGCTCCCTCTCCCACTCACTCAGT 752

Db	633	CCACGGTCCTGCGCCAGCCCAATAGATGTCCTGAGGCGCTGCCCTCTCCACTTCAGT	574	source	1. 519
QY	753	TCCCAATCTAAATTTTACAGAGATTCTGTTTGGGGAACTTAAGTCAGATCCAGAAC	812	/organism="Homo sapiens"	
Db	573	TCCCAATCTAAATTTTACAGAGATTCTGTTTGGGGAACTTAAGTCAGATCCAGAAC	514	/mol_type="mRNA"	
QY	813	CTTGGCTGCAAGGAGTCTGGGAAATGTCATTTCCCTAGAGGAAGTTAGGCTGGGTGGA	872	/db_xref="taxon:9606"	
Db	513	CTTGGCTGCAAGGAGTCTGGGAAATGTCATTTCCCTAGAGGAAGTTAGGCTGGGTGGA	454	/clone="DKFp762F1316"	
QY	873	GCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCAATCGTGAAGAACTCGGGAGAGGT	932	/tissue_type="melanoma (MeWo cell line)"	
Db	453	GCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCAATCGTGAAGAACTCGGGAGAGGT	394	/dev_stage="adult"	
QY	933	GGAGTCCACATCTAGGGTTGTCGCGCCCTGGCTCTATCCCTGCCAGAGGTGGGAAC	992	/lab_host="DH10B"	
Db	393	GGAGTCCACATCTAGGGTTGTCGCGCCCTGGCTCTATCCCTGCCAGAGGTGGGAAC	334	/clone_lib="762 (synonym: hmcl2)"	
QY	993	GGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCCTCCCGGCTTGACTTTCTTTCTAG	1052	/note="Vector: pSport1, Site_1: NotI; Site_2: SalI"	
Db	333	GGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCCTCCCGGCTTGACTTTCTTTCTAG	274		
QY	1053	TCCTGGGCGCTAGATTCTGCACTTTGGGCTCTCTGACACAAACACATCCCAAGTAGCC	1112		
Db	273	TCCTGGGCGCTAGATTCTGCACTTTGGGCTCTCTGACACAAACACATCCCAAGTAGCC	214		
QY	1113	GGAAGAGCTAAACACAGGGGTTCTTAAATGGCTGCGCCCGCCAGCCGCGGCTCCCTTG	1172		
Db	213	GGAAGAGCTAAACACAGGGGTTCTTAAATGGCTGCGCCCGCCAGCCGCGGCTCCCTTG	154		
QY	1173	GGCAAAAGGAATGTGACGCCCTTACCCCAACCTTCAACTACAGAACTCGGGCCACCCCA	1232		
Db	153	GGCAAAAGGAATGTGACGCCCTTACCCCAACCTTCAACTACAGAACTCGGGCCACCCCA	94		
QY	1233	GCAGTATTTTATTTAAATGTGCGCAATTTATGAGTTATGATCAATTTGTTAAATT	1292		
Db	93	GCAGTATTTTATTTAAATGTGCGCAATTTATGAGTTATGATCAATTTGTTAAATT	34		
QY	1293	AAAGTTACAGATGTCAAAAAAATAAAAAA	1325		
Db	33	AAAGTTACAGATGTCAAAAAAATAAAAAA	1		
RESULT 9					
AL138362					
LOCUS					
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
COMMENT					
FEATURES					
Db	633	CCACGGTCCTGCGCCAGCCCAATAGATGTCCTGAGGCGCTGCCCTCTCCACTTCAGT	574	Query Match	39.1%; Score 519; DB 9; Length 519;
QY	753	TCCCAATCTAAATTTTACAGAGATTCTGTTTGGGGAACTTAAGTCAGATCCAGAAC	812	Best Local Similarity	100.0%; Pred. No. 2e-82;
Db	573	TCCCAATCTAAATTTTACAGAGATTCTGTTTGGGGAACTTAAGTCAGATCCAGAAC	514	Matches	519; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	813	CTTGGCTGCAAGGAGTCTGGGAAATGTCATTTCCCTAGAGGAAGTTAGGCTGGGTGGA	872		
Db	513	CTTGGCTGCAAGGAGTCTGGGAAATGTCATTTCCCTAGAGGAAGTTAGGCTGGGTGGA	454		
QY	873	GCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCAATCGTGAAGAACTCGGGAGAGGT	932		
Db	453	GCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCAATCGTGAAGAACTCGGGAGAGGT	394		
QY	933	GGAGTCCACATCTAGGGTTGTCGCGCCCTGGCTCTATCCCTGCCAGAGGTGGGAAC	992		
Db	393	GGAGTCCACATCTAGGGTTGTCGCGCCCTGGCTCTATCCCTGCCAGAGGTGGGAAC	334		
QY	993	GGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCCTCCCGGCTTGACTTTCTTTCTAG	1052		
Db	333	GGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCCTCCCGGCTTGACTTTCTTTCTAG	274		
QY	1053	TCCTGGGCGCTAGATTCTGCACTTTGGGCTCTCTGACACAAACACATCCCAAGTAGCC	1112		
Db	273	TCCTGGGCGCTAGATTCTGCACTTTGGGCTCTCTGACACAAACACATCCCAAGTAGCC	214		
QY	1113	GGAAGAGCTAAACACAGGGGTTCTTAAATGGCTGCGCCCGCCAGCCGCGGCTCCCTTG	1172		
Db	213	GGAAGAGCTAAACACAGGGGTTCTTAAATGGCTGCGCCCGCCAGCCGCGGCTCCCTTG	154		
QY	1173	GGCAAAAGGAATGTGACGCCCTTACCCCAACCTTCAACTACAGAACTCGGGCCACCCCA	1232		
Db	153	GGCAAAAGGAATGTGACGCCCTTACCCCAACCTTCAACTACAGAACTCGGGCCACCCCA	94		
QY	1233	GCAGTATTTTATTTAAATGTGCGCAATTTATGAGTTATGATCAATTTGTTAAATT	1292		
Db	93	GCAGTATTTTATTTAAATGTGCGCAATTTATGAGTTATGATCAATTTGTTAAATT	34		
QY	1293	AAAGTTACAGATGTCAAAAAAATAAAAAA	1325		
Db	33	AAAGTTACAGATGTCAAAAAAATAAAAAA	1		
RESULT 10					
BX283437					
LOCUS					
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
COMMENT					
FEATURES					
Db	633	CCACGGTCCTGCGCCAGCCCAATAGATGTCCTGAGGCGCTGCCCTCTCCACTTCAGT	574	Query Match	39.1%; Score 519; DB 9; Length 519;
QY	753	TCCCAATCTAAATTTTACAGAGATTCTGTTTGGGGAACTTAAGTCAGATCCAGAAC	812	Best Local Similarity	100.0%; Pred. No. 2e-82;
Db	573	TCCCAATCTAAATTTTACAGAGATTCTGTTTGGGGAACTTAAGTCAGATCCAGAAC	514	Matches	519; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	813	CTTGGCTGCAAGGAGTCTGGGAAATGTCATTTCCCTAGAGGAAGTTAGGCTGGGTGGA	872		
Db	513	CTTGGCTGCAAGGAGTCTGGGAAATGTCATTTCCCTAGAGGAAGTTAGGCTGGGTGGA	454		
QY	873	GCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCAATCGTGAAGAACTCGGGAGAGGT	932		
Db	453	GCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCAATCGTGAAGAACTCGGGAGAGGT	394		
QY	933	GGAGTCCACATCTAGGGTTGTCGCGCCCTGGCTCTATCCCTGCCAGAGGTGGGAAC	992		
Db	393	GGAGTCCACATCTAGGGTTGTCGCGCCCTGGCTCTATCCCTGCCAGAGGTGGGAAC	334		
QY	993	GGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCCTCCCGGCTTGACTTTCTTTCTAG	1052		
Db	333	GGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCCTCCCGGCTTGACTTTCTTTCTAG	274		
QY	1053	TCCTGGGCGCTAGATTCTGCACTTTGGGCTCTCTGACACAAACACATCCCAAGTAGCC	1112		
Db	273	TCCTGGGCGCTAGATTCTGCACTTTGGGCTCTCTGACACAAACACATCCCAAGTAGCC	214		
QY	1113	GGAAGAGCTAAACACAGGGGTTCTTAAATGGCTGCGCCCGCCAGCCGCGGCTCCCTTG	1172		
Db	213	GGAAGAGCTAAACACAGGGGTTCTTAAATGGCTGCGCCCGCCAGCCGCGGCTCCCTTG	154		
QY	1173	GGCAAAAGGAATGTGACGCCCTTACCCCAACCTTCAACTACAGAACTCGGGCCACCCCA	1232		
Db	153	GGCAAAAGGAATGTGACGCCCTTACCCCAACCTTCAACTACAGAACTCGGGCCACCCCA	94		
QY	1233	GCAGTATTTTATTTAAATGTGCGCAATTTATGAGTTATGATCAATTTGTTAAATT	1292		
Db	93	GCAGTATTTTATTTAAATGTGCGCAATTTATGAGTTATGATCAATTTGTTAAATT	34		
QY	1293	AAAGTTACAGATGTCAAAAAAATAAAAAA	1325		
Db	33	AAAGTTACAGATGTCAAAAAAATAAAAAA	1		
RESULT 10					
BX283437					
LOCUS					
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
COMMENT					
FEATURES					
Db	633	CCACGGTCCTGCGCCAGCCCAATAGATGTCCTGAGGCGCTGCCCTCTCCACTTCAGT	574	Query Match	39.1%; Score 519; DB 9; Length 519;
QY	753	TCCCAATCTAAATTTTACAGAGATTCTGTTTGGGGAACTTAAGTCAGATCCAGAAC	812	Best Local Similarity	100.0%; Pred. No. 2e-82;
Db	573	TCCCAATCTAAATTTTACAGAGATTCTGTTTGGGGAACTTAAGTCAGATCCAGAAC	514	Matches	519; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	813	CTTGGCTGCAAGGAGTCTGGGAAATGTCATTTCCCTAGAGGAAGTTAGGCTGGGTGGA	872		
Db	513	CTTGGCTGCAAGGAGTCTGGGAAATGTCATTTCCCTAGAGGAAGTTAGGCTGGGTGGA	454		
QY	873	GCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCAATCGTGAAGAACTCGGGAGAGGT	932		
Db	453	GCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCAATCGTGAAGAACTCGGGAGAGGT	394		
QY	933	GGAGTCCACATCTAGGGTTGTCGCGCCCTGGCTCTATCCCTGCCAGAGGTGGGAAC	992		
Db	393	GGAGTCCACATCTAGGGTTGTCGCGCCCTGGCTCTATCCCTGCCAGAGGTGGGAAC	334		
QY	993	GGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCCTCCCGGCTTGACTTTCTTTCTAG	1052		
Db	333	GGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCCTCCCGGCTTGACTTTCTTTCTAG	274		
QY	1053	TCCTGGGCGCTAGATTCTGCACTTTGGGCTCTCTGACACAAACACATCCCAAGTAGCC	1112		
Db	273	TCCTGGGCGCTAGATTCTGCACTTTGGGCTCTCTGACACAAACACATCCCAAGTAGCC	214		
QY	1113	GGAAGAGCTAAACACAGGGGTTCTTAAATGGCTGCGCCCGCCAGCCGCGGCTCCCTTG	1172		
Db	213	GGAAGAGCTAAACACAGGGGTTCTTAAATGGCTGCGCCCGCCAGCCGCGGCTCCCTTG	154		
QY	1173	GGCAAAAGGAATGTGACGCCCTTACCCCAACCTTCAACTACAGAACTCGGGCCACCCCA	1232		
Db	153	GGCAAAAGGAATGTGACGCCCTTACCCCAACCTTCAACTACAGAACTCGGGCCACCCCA	94		
QY	1233	GCAGTATTTTATTTAAATGTGCGCAATTTATGAGTTATGATCAATTTGTTAAATT	1292		
Db	93	GCAGTATTTTATTTAAATGTGCGCAATTTATGAGTTATGATCAATTTGTTAAATT	34		
QY	1293	AAAGTTACAGATGTCAAAAAAATAAAAAA	1325		
Db	33	AAAGTTACAGATGTCAAAAAAATAAAAAA	1		
RESULT 10					
BX283437					
LOCUS					
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
COMMENT					
FEATURES					
Db	633	CCACGGTCCTGCGCCAGCCCAATAGATGTCCTGAGGCGCTGCCCTCTCCACTTCAGT	574	Query Match	39.1%; Score 519; DB 9; Length 519;
QY	753	TCCCAATCTAAATTTTACAGAGATTCTGTTTGGGGAACTTAAGTCAGATCCAGAAC	812	Best Local Similarity	100.0%; Pred. No. 2e-82;
Db	573	TCCCAATCTAAATTTTACAGAGATTCTGTTTGGGGAACTTAAGTCAGATCCAGAAC	514	Matches	519; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	813	CTTGGCTGCAAGGAGTCTGGGAAATGTCATTTCCCTAGAGGAAGTTAGGCTGGGTGGA	872		
Db	513	CTTGGCTGCAAGGAGTCTGGGAAATGTCATTTCCCTAGAGGAAGTTAGGCTGGGTGGA	454		
QY	873	GCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCAATCGTGAAGAACTCGGGAGAGGT	932		
Db	453	GCAAGCCCCACCTGCGTTTTTCTGCCACAGCATCAATCGTGAAGAACTCGGGAGAGGT	394		
QY	933	GGAGTCCACATCTAGGGTTGTCGCGCCCTGGCTCTATCCCTGCCAGAGGTGGGAAC	992		
Db	393	GGAGTCCACATCTAGGGTTGTCGCGCCCTGGCTCTATCCCTGCCAGAGGTGGGAAC	334		
QY	993	GGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCCTCCCGGCTTGACTTTCTTTCTAG	1052		
Db	333	GGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCCTCCCGGCTTGACTTTCTTTCTAG	274		
QY	1053	TCCTGGGCGCTAGATTCTGCACTTTGGGCTCTCTGACACAAACACATCCCAAGTAGCC	1112		
Db	273	TCCTGGGCGCTAGATTCTGCACTTTGGGCTCTCTGACACAAACACATCCCAAGTAGCC	214		
QY	1113	GGAAGAGCTAAACACAGGGGTTCTTAAATGGCTGCGCCCGCCAGCCGCGGCTCCCTTG	1172		
Db	213	GGAAGAGCTAAACACAGGGGTTCTTAAATGGCTGCGCCCGCCAGCCGCGGCTCCCTTG	154		
QY	1173	GGCAAAAGGAATGTGACGCCCTTACCCCAACCTTCAACTACAGAACTCGGGCCACCCCA	1232		
Db	153	GGCAAAAGGAATGTGACGCCCTTACCCCAACCTTCAACTACAGAACTCGGGCCACCCCA	94		
QY	1233	GCAGTATTTTATTTAAATGTGCGCAATTTATGAGTTATGATCAATTTGTTAAATT	1292		
Db	93	GCAGTATTTTATTTAAATGTGCGCAATTTATGAGTTATGATCAATTTGTTAAATT	34		
QY	1293	AAAGTTACAGATGTCAAAAAAATAAAAAA	1325		
Db	33	AAAGTTACAGATGTCAAAAAAATAAAAAA	1		
RESULT 10					
BX283437					
LOCUS					
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
COMMENT					
FEATURES					
Db	633	CCACGGTCCTGCGCCAGCCCAATAGATGTCCTGAGGCGCTGCCCTCTCCACTTCAGT	574	Query Match	39.1%; Score 519; DB 9; Length 519;
QY	753	TCCCAATCTAAATTTTACAGAGATTCTGTTTGGGGAACTTAAGTCAGATCCAGAAC	812	Best Local Similarity	100.0%; Pred. No. 2e-82;
Db	573	TCCCAATCTAAATTTTACAGAGATTCTGTTTGGGGAACTTAAGTCAGATCCAGAAC	514	Matches	519; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	813	CTTGGCTGCAAGGAGTCTGGGAAATGTCATTTCCCTAGAGGAAGTTAGGCTGGGTGGA	872		
Db	513	CTTGGCTGCAAGGAGTCTGGGAAATGTCATTTCCCTAGAGGAAGTTAGGCTGGGTGGA	454		

RZPD; IMAGE95805802.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human UnigeneSet - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-bin/showlib.pl.cgi/response?libNo=972 Contact: Ina Rolfes
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de

This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
PCMV-M13u, Primer sequence: CGTTGTAACGACGCCAGT.

FEATURES

Location/Qualifiers
1..510
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE95805802 ; IMAGE:3944860"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_7"
/note="Organ: lung; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 37.4%; Score 496; DB 13; Length 510;
Best Local Similarity 100.0%; Pred. No. 2.2e-78;
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCCACAGTGGTACAGGTAGTACCTGGTCTAGGGTTGCTGAGAGCCAACTCTCTCTGC 60
Db 15 TCCACAGTGGTACAGGTAGTACCTGGTCTAGGGTTGCTGAGAGCCAACTCTCTCTGC 74
QY 61 CACCCCCACACCAAGAACTATATGGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Db 75 CACCCCCACACCAAGAACTATATGGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 134
QY 121 GATGCTGTGCTGTGGAGGACCTGGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 180
Db 135 GATGCTGTGCTGTGGAGGACCTGGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 194
QY 181 CCACCTTCTCTGCCACAGGCGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
Db 195 CCACCTTCTCTGCCACAGGCGGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 254
QY 241 CCATTTAGCTTAAAGCACTGAGGACAAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Db 255 CCATTTAGCTTAAAGCACTGAGGACAAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 314
QY 301 CTCTGACAGAGTCCAGCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
Db 315 CTCTGACAGAGTCCAGCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 374
QY 361 GGAGGCTTCCAGAGCAGTGTGTATTTAGGACCCCAAGCACTGGGAGGGGCTGTGGCT 420
Db 375 GGAGGCTTCCAGAGCAGTGTGTATTTAGGACCCCAAGCACTGGGAGGGGCTGTGGCT 434
QY 421 AGACCCCTTGTGAGACTTGGCATCTATCTCAGTGTAGGATCTCTGCTGCAAAACAGAGC 480
Db 435 AGACCCCTTGTGAGACTTGGCATCTATCTCAGTGTAGGATCTCTGCTGCAAAACAGAGC 494
QY 481 CACTTGTAGCTGGTTT 496
Db 495 CACTTGTAGCTGGTTT 510

RESULT 11

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

ALL138363 501 bp mRNA linear EST 04-SEP-2003
DKFZp762F1316.s1.762 (synonym: hmcl2) Homo sapiens cDNA clone
DKFZp762F1316.3', mRNA sequence.
ALL138363
ALL138363.1 GI:6855044
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 501)
Ansorge,W., Wirkner,U., Mewes,W., Weil,B. and Wiemann,S.
EST (Ansorge,W., Wirkner,U., Mewes,H.W., Weil,B. and Wiemann,S.)
Unpublished (1999)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
r1 sequence also available.
This clone (DKFZp762F1316) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

Location/Qualifiers
1..501
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp762F1316"
/tissue_type="melanoma (MeWo cell line)"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="762 (synonym: hmcl2)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 36.9%; Score 489; DB 9; Length 501;
Best Local Similarity 100.0%; Pred. No. 3.7e-77;
Matches 489; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 820 GCAAGGGAGTCTGGAAATGTCATTTCCCTAGAAGGAAGTGGGTGGTGGGCAAGCC 879
Db 501 GCAAGGGAGTCTGGAAATGTCATTTCCCTAGAAGGAAGTGGGTGGTGGGCAAGCC 442
QY 880 CCACCTGGCTTTTCTGCCACAGCATCCCAATCGTGAAGAACTCGGAGAGGGTGGAGTCC 939
Db 441 CCACCTGGCTTTTCTGCCACAGCATCCCAATCGTGAAGAACTCGGAGAGGGTGGAGTCC 382
QY 940 ACATCTAGGGTGTCTGCGCCCTTGGCTCTATCCCTGCCAGAGGTGGAACTGGAGGAG 999
Db 381 ACATCTAGGGTGTCTGCGCCCTTGGCTCTATCCCTGCCAGAGGTGGAACTGGAGGAG 322
QY 1000 TGGGCTGCAAGACTCAGCTAAATGCTCTCCCGGGCTTGACTTTTCTTTCTAGTCTCTGG 1059
Db 321 TGGGCTGCAAGACTCAGCTAAATGCTCTCCCGGGCTTGACTTTTCTTTCTAGTCTCTGG 262
QY 1060 GCCTAGATTCTGCACCTTGGGGTCTCTGACAAACACACCATCCCAAGTAGCCGGAAG 1119
Db 261 GCCTAGATTCTGCACCTTGGGGTCTCTGACAAACACACCATCCCAAGTAGCCGGAAG 202
QY 1120 CTAACACAGGGGGTCTTAAATGGCTGCCGCCGCCACCCGGGCTCCCTTGGGCAAAA 1179
Db 201 CTAACACAGGGGGTCTTAAATGGCTGCCGCCGCCACCCGGGCTCCCTTGGGCAAAA 142
QY 1180 GGAATTGTGACCCCTACCCCAACCTTCACTACAGAACTCTGGCCACCCAGAGTAT 1239

```

Db      141  GGAATTGTGAGCCCTACCCCAACCCCTTCAACTACCAGAAATCTGGCCACCCAGCAGTAT 82
QY      1240  TTTTATTTAAATGTTGGCCCATTTTATGAGTTATGATCAATTTGTATTAATTAAGTTA 1299
Db      81    TTTTATTTAAATGTTGGCCCATTTTATGAGTTATGATCAATTTGTATTAATTAAGTTA 22
QY      1300  CAGATGTCA 1308
Db      21    CAGATGTCA 13

RESULT 12
AW515028/c
LOCUS   AW515028              470 bp      mRNA      linear      EST 03-MAR-2000
DEFINITION
x90805.x1 NCI_CGAP_Ut2 Homo sapiens cDNA clone IMAGE:2808992 3',
mRNA sequence.
ACCESSION
AW515028
VERSION  AW515028.1 GI:7153110
KEYWORDS
EST.
SOURCE   Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 470)
AUTHORS  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE    National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL  Unpublished (1997)
COMMENT  Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml
Seq primer: -40UP from Gibco
High quality sequence stop: 427.
FEATURES             source
    1..470
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:2808992"
        /tissue_type="moderately-differentiated endometrial
adenocarcinoma, 3 pooled tumors"
        /lab_host="DH10B"
        /clone_lib="NCI_CGAP_Ut2"
        /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.85 kb. Life Technologies catalog #:
11539-012"

ORIGIN
Query Match      34.3%; Score 455; DB 10; Length 470;
Best Local Similarity 100.0%; Pred. No. 3.6e-71;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      849  TAGAAGCAAGTAGGGTGGGTGGAGCAAGCCCACTGGCTTTTTCGCCACAGCATCCA 908
Db      455  TAGAAGCAAGTAGGGTGGGTGGAGCAAGCCCACTGGCTTTTTCGCCACAGCATCCA 396
QY      909  ATCTGTGAAGACTCGGGAGAGGGTGGAGTCCACATCTAGGGTGTGTCTGCCCTTGGCTC 969
Db      395  ATCTGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTGTGTCTGCCCTTGGCTC 336
QY      969  TATCCCTCCAGAGCTGGGAACTGGAGAGTGGGCTGCAGAGCTGAGCCTAAATGTCTC 1028
Db      335  TATCCCTGCCAGAGGTGGGAACTGGAGAGTGGGCTGCAGAGCTGAGCCTAAATGTCTC 276
QY      1029  CCGGGCCTTGACTTTCTTTCTAGTCTCTGGGGCCCTAGATTCTGCATTTGGGGTCTCTGAC 1088

```

```

Db      275  CCGGGCCTTGACTTTTCTTTCTAGTCTCTGGGCTAGATTCTGCATTTGGGGTCTCTGAC 216
QY      1089  ACAACACACATCCCAAGTAGCCGAAGAGCTAAACACAGGGGGTCTTAAATGGCTG 1148
Db      215  ACAACACACATCCCAAGTAGCCGAAGAGCTAAACACAGGGGGTCTTAAATGGCTG 156
QY      1149  CCCCCGCCACCCGGGCTCCCTTGGCCAAAGAAATTGTCCAGCCCTACCCCAACCCCTTCA 1208
Db      155  CCCCCGCCACCCGGGCTCCCTTGGCCAAAGAAATTGTCCAGCCCTACCCCAACCCCTTCA 96
QY      1209  ACTACCAGAATCTGGGCCACCCCGACGAGTATTTTATTTAAATGTTGCCATTTATGA 1268
Db      95    ACTACCAGAATCTGGGCCACCCCGACGAGTATTTTATTTAAATGTTGCCATTTATGA 36
QY      1269  GTTATGATCAATTTGTATTAATTAAGTTACAGA 1303
Db      35    GTTATGATCAATTTGTATTAATTAAGTTACAGA 1

RESULT 13
AI457621/c
LOCUS   AI457621              486 bp      mRNA      linear      EST 13-APR-1999
DEFINITION
tj63d11.x1 Soares NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
IMAGE:2146197 3', similar to contains Alu repetitive element; , mRNA
sequence.
ACCESSION
AI457621
VERSION  AI457621.1 GI:4310490
KEYWORDS
EST.
SOURCE   Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 486)
AUTHORS  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE    National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL  Unpublished (1997)
COMMENT  Contact: Robert Strausberg, Ph.D.
Email: cgapbs-rc@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 748 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 453.
FEATURES             source
    1..486
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:2146197"
        /lab_host="DH10B"
        /clone_lib="Soares NSF_F8_9W_OT_PA_P_S1"
        /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was from
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and cloneIDs: Soares NBHSF pool 1:
309384-310919, 323208-325895 Soares NBHSF pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares NBHSF-9W pool 1:
758280-760583, 772104-774407 Soares NBHSF pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NBHSF
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN
Query Match      33.2%; Score 440; DB 9; Length 486;
Best Local Similarity 100.0%; Pred. No. 1.5e-68;

```

Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 864 GTGGTGGAGCAAGCCCACTGGTGTCTTCTGCCAGCATCCAACTCGTGAAGAACTCG 923

Db 441 GTGGTGGAGCAAGCCCACTGGTGTCTTCTGCCAGCATCCAACTCGTGAAGAACTCG 382

QY 924 GGAGAGGTGGAGTCCACATCTAGGGTGTCTCTGCCCTTGGCTCTATCCCTGCCAGAG 983

Db 381 GGAGAGGTGGAGTCCACATCTAGGGTGTCTCTGCCCTTGGCTCTATCCCTGCCAGAG 322

QY 984 GTGGAACTGGAGAGTGGGCTGCAAGACTGAGCTTAAATGTCTCCCGGCTTGACTTT 1043

Db 321 GTGGAACTGGAGAGTGGGCTGCAAGACTGAGCTTAAATGTCTCCCGGCTTGACTTT 262

QY 1044 TCTTTCTAGTCTGGGCGCTAGATTCTGCATCTGGGCTCTCTGACACACACACCATCCC 1103

Db 261 TCTTTCTAGTCTGGGCGCTAGATTCTGCATCTGGGCTCTCTGACACACACACCATCCC 202

QY 1104 AAGTAGCCGAGAGCTAAACACAGGGGTCTTAAATGGTGTCCCGGCCACCCGGG 1163

Db 201 AAGTAGCCGAGAGCTAAACACAGGGGTCTTAAATGGTGTCCCGGCCACCCGGG 142

QY 1164 CTTCCCTTGGCAAAAGGAATGTGAGCTTACCCCAACCCCTTCACTACAGAACTGG 1223

Db 141 CTTCCCTTGGCAAAAGGAATGTGAGCTTACCCCAACCCCTTCACTACAGAACTGG 82

QY 1224 GCCACCCAGCAGTATTTTATTTAAATGTGCGCATTTTATGAGTTATGATCAATTG 1283

Db 81 GCCACCCAGCAGTATTTTATTTAAATGTGCGCATTTTATGAGTTATGATCAATTG 22

QY 1284 TATTAATTAAGTTACAGA 1303

Db 21 TATTAATTAAGTTACAGA 2

RESULT 14

CA434008/c

LOCUS

DEFINITION

UI-H-DF0-arq-p-13-0-UI.s1 NCI CGAP DF0 Homo sapiens cDNA clone

UI-H-DF0-arq-p-13-0-UI 3', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 451)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Dr. Jose Mercuende

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu

Seq primer: M13 FORWARD

POLYA=Yes.

FEATURES

source

1. 451

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-H-DF0-arq-p-13-0-UI"

/tissue_type="Subchondral Bone"

/dev_stage="Adult"

/lab_host="DH10B (Life Technologies)"

/clone_lib="NCI CGAP DF0"

/note="Organ: Bone; Vector: pT7T3-Pac (Pharmacia) with a

ORIGIN

Query Match 32.7%; Score 434; DB 14; Length 451;

Best Local Similarity 100.0%; Pred. No. 1.8e-67;

Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 870 GGAGCAAGCCCACTGGTGTCTTCTGCCAGCATCCAACTCGTGAAGAACTCGGAGAG 929

Db 451 GGAGCAAGCCCACTGGTGTCTTCTGCCAGCATCCAACTCGTGAAGAACTCGGAGAG 392

QY 930 GGTGGAGTCCACATCTAGGGTGTCTCTGCCCTTGGCTCTATCCCTGCCAGAGTGGGA 989

Db 391 GGTGGAGTCCACATCTAGGGTGTCTCTGCCCTTGGCTCTATCCCTGCCAGAGTGGGA 332

QY 990 ACTGGAGGAGTGGGCTGCAAGCTGAGCCTAAATGTCTCCCGGCTTGACTTTCTTTC 1049

Db 331 ACTGGAGGAGTGGGCTGCAAGCTGAGCCTAAATGTCTCCCGGCTTGACTTTCTTTC 272

QY 1050 TAGTCTCTGGGCGCTAGATTCTGCACCTTGGGCTCTCTGACACACACACCATCCAAAGTA 1109

Db 271 TAGTCTCTGGGCGCTAGATTCTGCACCTTGGGCTCTCTGACACACACACCATCCAAAGTA 212

QY 1110 GCGGGAAGAGCTAAACACAGGGGTCTTAAATGGCTGCCCGGCCACCCGGGCTCCC 1169

Db 211 GCGGGAAGAGCTAAACACAGGGGTCTTAAATGGCTGCCCGGCCACCCGGGCTCCC 152

QY 1170 TTGGGCAAAAGAAATGTGAGCCTACCCCAACCCCTTCACTACCAAGATCTGGGCCACC 1229

Db 151 TTGGGCAAAAGAAATGTGAGCCTACCCCAACCCCTTCACTACCAAGATCTGGGCCACC 92

QY 1230 CCAGCAGTATTTTATTTAAATGTGCGCATTTTATGAGTTATGATCAATTGTATTAA 1289

Db 91 CCAGCAGTATTTTATTTAAATGTGCGCATTTTATGAGTTATGATCAATTGTATTAA 32

QY 1290 ATTAAGTTACAGA 1303

Db 31 ATTAAGTTACAGA 18

RESULT 15

BE671191/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BE671191

7e55b02.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:3286347 3', mRNA sequence.

477 bp mRNA linear EST 08-SEP-2000

BE671191

BE671191.1 GI:10031732

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 477)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov


```

RESULT 17
AI066431/c
LOCUS
DEFINITION
  oq98e10.x1 NCI_CGAP_Col2 Homo sapiens cDNA clone IMAGE:1594410 3',
  mRNA sequence.
ACCESSION
  AI066431
VERSION
  AI066431.1 GI:3367133
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgaps-remail.nih.gov
  Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
  Emmert-Buck, M.D., Ph.D.
  cDNA Library Preparation: Stratagene, Inc.
  DNA sequencing by: Greg Lennon, Ph.D.
  Cloning distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bhrp/image/image.html
  Insert length: 1019 Std Error: 0.00
  Seq primer: -40m13 fwd. ET from Amersham
  High quality sequence stop: 374.
  Location/Qualifiers
    1..423
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:1594410"
      /sex="mixed"
      /tissue_type="colon tumor"
      /lab_host="SOLR (kanamycin resistant)"
      /clone_lib="NCI_CGAP_Col2"
      /note="Organ: colon; Vector: Bluescript SK-; Site 1:
      EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
      Oligo dT. Pooled colon tumors. 5' adaptor sequence: 5'
      GAATCGGCACGAG 3' 3' adaptor sequence: 5'
      CTCAGTGTGTTTTTTTTTTT 3' Average insert size: 1.2 kb."
ORIGIN
  Query Match 31.7%; Score 421; DB 9; Length 423;
  Best Local Similarity 100.0%; Pred. No. 3.6e-65;
  Matches 421; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 888 GTTTTCTGCCACAGCATCAATCGTAGAAGTCTGGGAGAGGTGGATCCACATCTAG 947
  Db 423 GTTTTCTGCCACAGCATCAATCGTAGAAGTCTGGGAGAGGTGGATCCACATCTAG 364
  QY 948 GGTGTCTCTGCCCTTGGCTCTATCCTCGCAGAGTGGGAAGTGGAGAGTGGGCTGC 1007
  Db 363 GGTGTCTCTGCCCTTGGCTCTATCCTCGCAGAGTGGGAAGTGGAGAGTGGGCTGC 304
  QY 1008 AAGACTGAGCCTAAATGTCTCCCGGCTTGTCTTCTTCTAGTCCTGGGCTTAGAT 1067
  Db 303 AAGACTGAGCCTAAATGTCTCCCGGCTTGTCTTCTTCTAGTCCTGGGCTTAGAT 244
  QY 1068 TCTGCACTTGGGCTCTCTGACACACACACACACACACACACACACACACACACAC 1127
  Db 243 TCTGCACTTGGGCTCTCTGACACACACACACACACACACACACACACACACACAC 184
  QY 1128 AGGGGTTCTTAAATGGTGTCCCGGCTCCCGGCTCCCTTGGGCAAAAGGAATGT 1187
  Db 183 AGGGGTTCTTAAATGGTGTCCCGGCTCCCGGCTCCCTTGGGCAAAAGGAATGT 124

```

```

QY 1188 CAGCCTACCCCAACCCCTTCACTACCAAGATCTGGGCCACCCAGAGTATTTTATT 1247
Db 123 CAGCCTACCCCAACCCCTTCACTACCAAGATCTGGGCCACCCAGAGTATTTTATT 64
QY 1248 AAAATGTTGCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAGTTACAGATGC 1307
Db 63 AAAATGTTGCCATTTTATGAGTTATGATCAATTTGTATTAAATTAAGTTACAGATGC 4
QY 1308 A 1308
Db 3 A 3
  BQ187235 501 bp mRNA linear EST 30-APR-2002
  UI-E-EJ1-ajz-i-06-0-UI.r1 UI-E-EJ1 Homo sapiens cDNA clone
  DEFINITION
  BQ187235 UI-E-EJ1-ajz-i-06-0-UI 5', mRNA sequence.
  ACCESSION
  BQ187235
  VERSION
  BQ187235.1 GI:20362786
  KEYWORDS
  EST.
  SOURCE
  Homo sapiens (human)
  ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
  1 (bases 1 to 501)
  AUTHORS
  Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  JOURNAL
  Genome Res. 6 (9), 791-806 (1996)
  MEDLINE
  97044477
  PUBMED
  889548
  COMMENT
  Contact: Soares, MB
  Coordinated Laboratory for Computational Genomics
  University of Iowa
  375 Newton Road, 4156 MEBRP, Iowa City, IA 52242, USA
  Tel: 319 335 8250
  Fax: 319 335 9565
  Email: bento-soares@uiowa.edu
  Tissue Procurement: Dr. Gregg Hageman
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  DNA sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Researchers may obtain clones from Research
  Genetics (www.resgen.com).
  The following repetitive elements were found in this cDNA
  sequence: 470-496. >POLY_A$imple_repeat
  Seq primer: M13 REVERSE
  Location/Qualifiers
    1..501
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="UI-E-EJ1-ajz-i-06-0-UI"
      /tissue_type="fetal eyes, lens, eye anterior segment,
      optic nerve, retina, Retina Foveal and Macular, RPE and
      Choroid"
      /dev_stage="fetal and adult"
      /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
      /clone_lib="UI-E-EJ1"
      /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
      modified polylinker; Site 1: EcoR I; Site 2: Not I;
      UI-E-EJ1 is a subtracted cDNA library constructed
      according to Bonaldo, Lennon and Soares, Genome Research,
      6:791-806, 1996. First strand cDNA synthesis was primed
      with an oligo-dT primer containing a Not I site. Double
      stranded cDNA was ligated to an EcoR I adaptor, digested
      with Not I, and cloned directionally into pT73-Pac
      vector. The oligonucleotide used to prime the synthesis of
      first-strand cDNA contains a library tag sequence that is
      located between the Not I site and the (dT)18 tail. The
      sequence tags for this library are: fetal eyes,
      AGAATCAAGA; lens, CGATTAGCGA; eye anterior segment,

```

AATGGCGCAT; optic nerve, CCATTAAGTG; retina, CCGCG; Retina
Foveal and Macular, GTCC; RPE and Choroid, ACTA. This
library was created for the program, Gene Discovery in the
Visual System, supported by National Eye Institute (NEI)."

ORIGIN

Query Match 31.7%; Score 421; DB 13; Length 501;
Best Local Similarity 99.8%; Pred. No. 3e-65; 1; Indels 0; Gaps 0;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 832 GGGAAATGTCATTTCCCTAGAGGAAGTTAGGCTGGGTGGAGCAAGCCACCTGGTGT 891
Db 1 GGGAAATGTCATTTCCCTAGAGGAAGTTAGGCTGGGTGGAGCAAGCCACCTGGTGT 60
QY 892 TTCTGCCACAGCATCCAAATCGTGAAGAACTCGGAGAGGCTGGAGTCCACATCTAGGTT 951
Db 61 TTCTGCCACAGCATCCAAATCGTGAAGAACTCGGAGAGGCTGGAGTCCACATCTAGGTT 120
QY 952 GTCTGCGCCCTGGCTTCATCCCTGCGAGAGGTGGAACTGGAGGAGTGGCTGCAAGA 1011
Db 121 GTCTGCGCCCTGGCTTCATCCCTGCGAGAGGTGGAACTGGAGGAGTGGCTGCAAGA 180
QY 1012 CTGAGCCTAAATGTCCTCCCGGCTTGACATTTTCTTTCTAGTCTGGGGCTAGATCTG 1071
Db 181 CTGAGCCTAAATGTCCTCCCGGCTTGACATTTTCTTTCTAGTCTGGGGCTAGATCTG 240
QY 1072 CACTTGGGGTCTTGACACAAACACCATCCAAAGTACCGGGAAGAGCTAAACACAGG 1131
Db 241 CACTTGGGGTCTTGACACAAACACCATCCAAAGTACCGGGAAGAGCTAAACACAGG 300
QY 1132 GGTTCCTAAATGCTGCGCCGCTCCCTGGGCTCCCTGGGCAAGGAATGTGAGC 1191
Db 301 GGTTCCTAAATGCTGCGCCGCTCCCTGGGCTCCCTGGGCAAGGAATGTGAGC 360
QY 1192 CCTACCCCAACCTTCAACTACCAATCTGGGCAACCCAGCAGTATTTTATTTAAA 1251
Db 361 CCTACCCCAACCTTCAACTACCAATCTGGGCAACCCAGCAGTATTTTATTTAAA 420
QY 1252 TGTTCGCCATTTATGAGTTATGATCAATTTGATTAATTAAGTTACAGA 1303
Db 421 TGTTCGCCATTTATGAGTTATGATCAATTTGATTAATTAAGTTACAGA 472

RESULT 19
BE795637 983 bp mRNA linear EST 20-SEP-2000
LOCUS 601590620F1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:3944860 5',
DEFINITION mRNA sequence.
ACCESSION BE795637
VERSION BE795637.1 GI:10215835
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 983)
NIH-MGC http://www.ncbi.nlm.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTP/DTF
cDNA library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLC802 row: 9 column: 05
High quality sequence stop: 791.
Location/Qualifiers
1. .983
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES

source

/db_xref="taxon:9606"
/clone="IMAGE:3944860"
/tissue_type="small cell carcinoma"
/cell_line="WGC3"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 7"
/note="Organ: lung; Vector: pOTB7; Site: 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 31.3%; Score 415; DB 10; Length 983;
Best Local Similarity 99.8%; Pred. No. 1.8e-64;
Matches 535; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 9 GGTACAGGTAGTACCTGTCTAGGTTTCCTTGAGAGCCAACTCTCTGSCCACCACCCCA 68
Db 1 GGTACAGGTAGTACCTGTCTAGGTTTCCTTGAGAGCCAACTCTCTGSCCACCACCCCA 60
QY 69 CACCAAGAACTATATGTTCTCTACTCTCCACTGATCTGCTGATGATGATGCTGT 128
Db 61 CACCAAGAACTATATGTTCTCTACTCTCCACTGATCTGCTGATGATGATGCTGT 120
QY 129 GGCTGTGGAAGGCACTGTGATGTCACACATTATAGTCCACCATCTTCACACCATCTC 188
Db 121 GGCTGTGGAAGGCACTGTGATGTCACACATTATAGTCCACCATCTTCACACCATCTC 180
QY 189 CTGCCCCACAGCCGAGGACAGGCTGAGGTATATCCCAAGCTGATGACAGGCCCATTTAG 248
Db 181 CTGCCCCACAGCCGAGGACAGGCTGAGGTATATCCCAAGCTGATGACAGGCCCATTTAG 240
QY 249 CCTAAAAGCAACTGACAGGACAAAGCTCCCTGGATGATCGAGGTCCCAAGTCTGTAAC 308
Db 241 CCTAAAAGCAACTGACAGGACAAAGCTCCCTGGATGATCGAGGTCCCAAGTCTGTAAC 300
QY 309 AAGAGTCCAGCAACCTCTTCAGCCAGGCTCTGTGACCTGTAGGTGTCAGAGGCTT 368
Db 301 AAGAGTCCAGCAACCTCTCTTCAGCCAGGCTCTGTGACCTGTAGGTGTCAGAGGCTT 360
QY 369 CCAGAAGCAGTGTGTAAATTAGGACCAAGCACTGGGAGGGCTGTGGCTAGACCCCT 428
Db 361 CCAGAAGCAGTGTGTAAATTAGGACCAAGCACTGGGAGGGCTGTGGCTAGACCCCT 419
QY 429 TGTGAGACTTGGCATCTATCTCAGTTAGGATCTCTGCTCCAGAAAACAGAGCCACTGTA 488
Db 420 TGTGAGACTTGGCATCTATCTCAGTTAGGATCTCTGCTCCAGAAAACAGAGCCACTGTA 479
QY 489 GCTGGTTTAAATAGACAGGATTTACTACCTGGCCCTGTGGCTTGCATAATTGT 544
Db 480 GCTGGTTTAAATAGACAGGATTTACTACCTGGCCCTGTGGCTTGCATAATTGT 535

RESULT 20

BE856182/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

BE856182 456 bp mRNA linear EST 29-SEP-2000
7f89g01.x1 NCI_CGAP_Pr28 Homo sapiens cDNA clone IMAGE:3304176 3',
similar to contains Alu repetitive element; mRNA sequence.

BE856182

BE856182.1 GI:10368957

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 456)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL
COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapb-re@mail.nih.gov
Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/INLNL, send email to: info@image.llnl.gov

Seq primer: -40UP from Gibco.

FEATURES
source

Location/Qualifiers
1..456
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3304176"
/sex="male"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="NCI CGAP Pr28"
/note="Organ: Prostate; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Pr22 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids) 95608-986759, 1101192-1101959, and 1217928-1220615. Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 30.5%; Score 405; DB 10; Length 456;
Best Local Similarity 99.8%; Pred. No. 2.1e-62;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 847 CCTAGAGGAAGTTAGGGTGGGAGCAAGCCACCTCGGTTTCTGCGCACACATC 906
Db
456 CCTAGAGGAAGTTAGGGTGGGAGCAAGCCACCTCGGTTTCTGCGCACATC 397
QY 907 CAATCGTGAAGAACTCGGAGAGGGTGGAGTCCACATCTAGGTTCTCTGCCCTTGGC 966
Db
396 CAATCGTGAAGAACTCGGAGAGGGTGGAGTCCACATCTAGGTTCTCTGCCCTTGGC 337
QY 967 TCTATCCCTGCCAGAGTGGGAATCGGAGGTGGCTGCAGACTGAGCCTAAATGTC 1026
Db
336 TCTATCCCTGCCAGAGTGGGAATCGGAGGTGGCTGCAGACTGAGCCTAAATGTC 277
QY 1027 TCCCGGCTTGACATTTCTTTCTAGTCTCGGGCCCTAGATTCTGCACCTGGGCTCTG 1086
Db
276 TCCCGGCTTGACATTTCTTTCTAGTCTCGGGCCCTAGATTCTGCACCTGGGCTCTG 217
QY 1087 ACAGAACACACATCCCAAGTAGCGGAAGAGCTAAACACAGGGGTTCTTAAATGGC 1146
Db
216 ACAGAACACACATCCCAAGTAGCGGAAGAGCTAAACACAGGGGTTCTTAAATGGC 157
QY 1147 TGCCCCGCCACCCGGGCTCCCTTGGGCAAGAGATTTGTGAGCCTACCCACCCCTT 1206
Db
156 TGCCCCGCCACCCGGGCTCCCTTGGGCAAGAGATTTGTGAGCCTACCCACCCCTT 97
QY 1207 CAATACCAGATCTGGGCCACCCAGCAGTATTTTATTTAAATGTTGCCATTTAT 1266
Db
96 CAATACCAGATCTGGGCCACCCAGCAGTATTTTATTTAAATGTTGCCATTTAT 37
QY 1267 GAGTTATGATCAATTTGATTAATTAAGTTACAG 1302
Db
36 GAGTTATGATCAATTTGATTAATTAAGTTACAG 1

RESULT 21

AA700789/c

LOCUS

AA700789 456 bp mRNA linear EST 19-DEC-1997

DEFINITION

z166c04.s1 Soares_fetal_liver_spleen_INFLS_S1 Homo sapiens cDNA clone IMAGE:435750 3', mRNA sequence.

ACCESSION

AA700789

VERSION

AA700789.1 GI:2703954

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 456)

AUTHORS

Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisels,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Scheilenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wyllie,T., Waterston,R. and Wilson,R. WashU-NCI human EST Project

JOURNAL

Unpublished (1997)

COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through INLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40ml3 fwd. E1 from Amersham

High quality sequence stop: 439.

FEATURES

source

Location/Qualifiers
1..456
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GBB:1335519"
/db_xref="taxon:9606"
/clone="IMAGE:435750"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="PH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS_S1"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Pac I; Site 2: Eco RI; This is a subtracted version of the original Soares fetal liver spleen INFLS library. 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5 RACTGAGAAATTAATTAAGATCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 30.4%; Score 403; DB 9; Length 456;
Best Local Similarity 99.8%; Pred. No. 4.8e-62;
Matches 453; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 855 GAAGTTAGGGTGGGAGAGTCCACATCTAGGGTCTGCTGCCCTTGGCTCTATCCC 914
Db
456 GAAGTTAGGGTGGGAGAGTCCACATCTAGGGTCTGCTGCCCTTGGCTCTATCCC 397
QY 915 AAGAACTCGGAGAGGGTGGAGTCCACATCTAGGGTCTGCTGCCCTTGGCTCTATCCC 974
Db
396 AAGAACTCGGAGAGGGTGGAGTCCACATCTAGGGTCTGCTGCCCTTGGCTCTATCCC 337
QY 975 TGCCAGAGGTGGGAACCTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCCTCCCGGC 1034
Db
336 TGCCAGAGGTGGGAACCTGGAGGAGTGGGCTGCAAGACTGAGCCTAAATGTCCTCCCGGC 277
QY 1035 CTTGACATTTCTTTCTAGTCTGGGGCCTAGATTCTGCATTTGGGCTCTGACACACA 1094
Db
276 CTTGACATTTCTTTCTAGTCTGGGGCCTAGATTCTGCATTTGGGCTCTGACACACA 217
QY 1095 CACCATCCCAAGTAGCGGAGAGCTAAACACAGGGGGTCTTAAATGCTGCCCGGC 1154
Db
216 CACCATCCCAAGTAGCGGAGAGCTAAACATAGGGGGTCTTAAATGCTGCCCGGC 157

sequence tags for this library are: fetal eyes, AATCCAGAT; lens, CGATTAGCA; eye anterior segment, AATCCAGAT; optic nerve, CCATTAGTG; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI). TAG TISSUE=RPE and Choroid TAG_LIB=UI-E-EJ1 TAG_SEQ=ACCTA"

1155 CCACCGGGGCTCCCTTGGGCAAAAGAAATGTGACGCTTACCCCAACCTTCAACTACC 1214
156 CCACCGGGGCTCCCTTGGGCAAAAGAAATGTGACGCTTACCCCAACCTTCAACTACC 97
1215 AGAATCTGGGCAACCCAGAGTATTTTATTTAAATGTTGCCATTTTATGAGTTATG 1274
96 AGAATCTGGGCAACCCAGAGTATTTTATTTAAATGTTGCCATTTTATGAGTTATG 37
1275 ATCAATTTGTTAAATTTAAAGTTACAGATGTCA 1308
36 ATCAATTTGTTAAATTTAAAGTTACAGATGTCA 3

RESULT 22
CK300567/c
LOCUS
DEFINITION
UI-E-EJ1-aj2-i-06-0-UI.s1 UI-E-EJ1 Homo sapiens cDNA clone
UI-E-EJ1-aj2-i-06-0-UI 3', mRNA sequence.
CK300567
CK300567.1 GI:39890062
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 502)
Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Greg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/eye.html
The following repetitive elements were found in this cDNA
sequence: 1-40, >AT rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.

Location/Qualifiers
1..502
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EJ1-aj2-i-06-0-UI"
/tissue_type="fetal eyes, lens, eye anterior segment,
optic nerve, retina, Retina Foveal and Macular, RPE and
Choroid"
/dev_stages="fetal and adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EJ1"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-EJ1 is a subtracted cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The

ORIGIN

Query Match 30.1%; Score 399; DB 14; Length 502;
Best Local Similarity 100.0%; Pred. No. 2.2e-61;
Matches 399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 832 GCGAAATGTCATTTCCTAGAGGAAGTATAGGCTGGGTGGAGCAAGCCACCTCGCTTT 891
Db 493 GCGAAATGTCATTTCCTAGAGGAAGTATAGGCTGGGTGGAGCAAGCCACCTCGCTTT 434
QY 892 TTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTT 951
Db 433 TTCTGCCACAGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTT 374
QY 952 GTCTGCGCCCTTGCTCTATCCCTGCCAGAGTGGAACTGGAGGAGTGGCTGCAAGA 1011
Db 373 GTCTGCGCCCTTGCTCTATCCCTGCCAGAGTGGAACTGGAGGAGTGGCTGCAAGA 314
QY 1012 CTGAGCCTAATGTCCTCCCGGCTTGACATTTCTTTCTAGTCTCTGGGGCTAGATTCTG 1071
Db 313 CTGAGCCTAATGTCCTCCCGGCTTGACATTTCTTTCTAGTCTCTGGGGCTAGATTCTG 254
QY 1072 CACTTGGGGTCTCTGACACAAACACCATCCCAAAGTAGCGGAAGAGCTAAACACAGGG 1131
Db 253 CACTTGGGGTCTCTGACACAAACACCATCCCAAAGTAGCGGAAGAGCTAAACACAGGG 194
QY 1132 GGTTCCTTAAATGCTGCCCGCCCGCCCGGCTCCCTTGGGCAAGGAATTGTGAGC 1191
Db 193 GGTTCCTTAAATGCTGCCCGCCCGCCCGGCTCCCTTGGGCAAGGAATTGTGAGC 134
QY 1192 CCTACCCCAACCTTCAACTACAGAAATCTGGGCCACCC 1230
Db 133 CCTACCCCAACCTTCAACTACAGAAATCTGGGCCACCC 95

RESULT 23
AI681374/c
LOCUS
DEFINITION
tx46a12.X1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2272606 3',
mRNA sequence.
ACCESSION
AI681374
VERSION
AI681374.1 GI:4891556
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 512)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 1330 Std Error: 0.00
Seq primer: -400P from Gibco

FEATURES	High quality sequence stop: 444.		REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
	Location/Qualifiers			1 (bases 1 to 431)	
	source			Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,	
	/organism="Homo sapiens"			Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,	
	/mol_type="mRNA"			Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,	
	/db_xref="taxon:9606"			Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P. and	
	/clone="IMAGE:2272606"			Wilson, R.	
	/tissue_type="carcinoid"			The WashU-Merck EST Project	
	/lab_host="DH10B"			Unpublished (1995)	
	/lab_host="NCI CGAP Lu24"			Contact: Wilson RK	
TITLE	/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a		JOURNAL	Washington University School of Medicine	
	modified polylinker; Plasmid DNA from the normalized			4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108	
	library NCI CGAP Lu5 was prepared, and ss circles were			Tel: 314 286 1800	
	made in vitro. Following HAP purification, this DNA was			Fax: 314 286 1810	
	used as tracer in a subtractive hybridization reaction.			Email: est@watson.wustl.edu	
	The driver was PCR-amplified cDNAs from a pool of 5,000			Insert Size: 855	
	clones made from the same library (clonoids			High quality sequence stops: 325	
	1414920-1417991 and 1520904-1522439). Subtraction by Bento			Source: IMAGE Consortium, LLNL	
	Soares and M. Fatima Ronaldo.			This clone is available royalty-free through LLNL; contact the	
				IMAGE Consortium (info@image.llnl.gov) for further information.	
ORIGIN	Query Match 29.5%; Score 391; DB 9; Length 512;		FEATURES	Location/Qualifiers	
	Best Local Similarity 99.8%; Pred. No. 5.4e-60;			1. .431	
	Matches 511; Conservative 0; Mismatches 0; Indels 1; Gaps 1;			/organism="Homo sapiens"	
				/mol_type="mRNA"	
				/db_xref="GDB:552113"	
				/db_xref="taxon:9606"	
				/clone="IMAGE:143000"	
				/sex="Female"	
				/dev_stage="placenta obtained at birth (full term)"	
				/lab_host="DH10B (ampicillin resistant)"	
QY	798 AGTCAGATCCAGAACCTTGGCTGCAGGAGTCTGGGAATGTCATTCCCTAGAGGAA 857		source	/clone_lib="Soares placenta Nb2HP"	
	512 AGTCAGATCCAGAACCTTGGCTGCAGGAGTCTGGGAATGTCATTCCCTAGAGGAA 453			/note="Organ: placenta; Vector: pT73D (Pharmacia) with a	
	858 GTTAGGGTGGGTGGAGCAAGCCCACTGCGTTTCTGCCAGACATCAATCGTGAAG 917			modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st	
	452 GTTAGGGTGGGTGGAGCAAGCCCACTGCGTTTCTGCCAGACATCAATCGTGAAG 393			strand cDNA was primed with a Not I - oligo(dT) primer [5',	
	918 AACTCGGGAGAGGTGGAGTCCACATCTAGGGTTGTCTGCCAGACATCAATCGTGAAG 977			AACTCGAAGATTTCGCGCGCAGAGATTTTTTTTTTTTTTTT 3'],	
	392 AACTCGGGAGAGGTGGAGTCCACATCTAGGGTTGTCTGCCAGACATCAATCGTGAAG 333			double-stranded cDNA was ligated to Eco RI adaptors	
	978 CCAGAGGTGGGAACTGGAGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCGGCCCT 1036			(Pharmacia), digested with Not I and cloned into the Not I	
	332 CCAGAGGTGGGAACTGGAGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCGGCCCT 273			and Eco RI sites of the modified pT73 vector. Library	
	1037 TGACTTTTCTTCTAGTCTCTGGGGCTAGATTCTGCACATTTGGGGTCTCTGACACACACA 1096			went through one round of normalization. Library	
	272 TGACTTTTCTTCTAGTCTCTGGGGCTAGATTCTGCACATTTGGGGTCTCTGACACACACA 213			constructed by Bento Soares and M.Fatima Ronaldo. "	
QY	1097 CCATCCAAAGTAGCCGGAAGAGCTAAACACAGGGGTCTTAAATGGTGTCCCGGCC 1156		ORIGIN	Query Match 28.5%; Score 378; DB 14; Length 431;	
	212 CCATCCAAAGTAGCCGGAAGAGCTAAACACAGGGGTCTTAAATGGTGTCCCGGCC 153			Best Local Similarity 100.0%; Pred. No. 1.2e-57;	
	1157 ACCCGGCTCTCTGGGCAAGAGGATTCGAGCCCTACCCCAACCCCTTCACTACCAG 1216			Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
	152 ACCCGGCTCTCTGGGCAAGAGGATTCGAGCCCTACCCCAACCCCTTCACTACCAG 93				
	1217 AATCTGGGCAACCCAGCAGTATTTTATTTAAATGTTGCCCATTTTATGAGTTATGAT 1276			QY 918 AACTCGGAGAGGGTGGAGTCCACATCTAGGGTTGTCTGCCCTTGGCTCTATCCCTGC 977	
	92 AATCTGGGCAACCCAGCAGTATTTTATTTAAATGTTGCCCATTTTATGAGTTATGAT 33			Db 391 AACTCGGAGAGGGTGGAGTCCACATCTAGGGTTGTCTGCCCTTGGCTCTATCCCTGC 332	
	1277 CAATTGTATTAATTAAGTTACAGATGCA 1308			QY 978 CCAGAGGTGGGAACTGGAGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCGGCCCT 1037	
	32 CAATTGTATTAATTAAGTTACAGATGCA 1			Db 331 CCAGAGGTGGGAACTGGAGAGTGGGCTGCAAGACTGAGCCTAAATGTCTCCCGGCCCT 272	
				QY 1038 GACTTTTCTTCTAGTCTCTGGGGCTAGATTCTGCACATTTGGGGTCTCTGACACACACAC 1097	
				Db 271 GACTTTTCTTCTAGTCTCTGGGGCTAGATTCTGCACATTTGGGGTCTCTGACACACACAC 212	
RESULT 24	R71654/c		QY	1098 CATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGTCTTAAATGGTGTCCCGGCCCA 1157	
	DEFINITION			211 CATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGTCTTAAATGGTGTCCCGGCCCA 152	
	Y153905.s1 Soares placenta Nb2HP Homo sapiens cDNA clone			1158 CCGGGGCTCTCTGGGCAAGAGGATTTGTGAGCCCTACCCCAACCCCTTCACTACCAG 1217	
	IMAGE:143000 3', mRNA sequence.			151 CCGGGGCTCTCTGGGCAAGAGGATTTGTGAGCCCTACCCCAACCCCTTCACTACCAG 92	
	R71654			1218 ATCTGGGCCACCCAGCAGTATTTTATTTAAATGTTGCCCATTTTATGAGTTATGATC 1277	
	R71654.1			91 ATCTGGGCCACCCAGCAGTATTTTATTTAAATGTTGCCCATTTTATGAGTTATGATC 32	
	GI:845171				
	EST.				
	Homo sapiens (human)				
	Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					

```

Qy 1278 AATTGTATTAAATATAA 1295
|||||
Db 31 AATTGTATTAAATATAA 14

RESULT 25
BE244684 497 bp mRNA linear EST 03-OCT-2001
LOCUS TCBAP2E0532 Pediatric pre-B cell acute lymphoblastic leukemia
DEFINITION Baylor-HGSC project=TCBA Homo sapiens cDNA clone TCBAP0532, mRNA
sequence.
ACCESSION BE244684
VERSION BE244684.1 GI:9096426
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 497)
AUTHORS Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman Jr., F.R.,
Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.,
TITLE Pediatric leukemia cDNA Sequencing Project
JOURNAL Unpublished (2000)
COMMENT Contact: Dr. Judith F. Margolin
Texas Children's Cancer Center and Human Genome Sequencing Center
at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536
Fax: 832-825-4038
Email: clones@txccc.org
Citation: Carninci, P. and Hayashizaki, Y. High efficiency
full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Seq primer: M13 primer.

FEATURES
source
Location/Qualifiers
1..497
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="TCBAP0532"
/sex="male"
/tissue_type="leukopheresis"
/cell_type="pre-B cell"
/dev_stage="pediatric 2 years"
/lab_host="DH10B"
/clone_lib="pediatric pre-B cell acute lymphoblastic
leukemia Baylor-HGSC project=TCBA"
/notes="vector: lambda pSB; Site.1: BamHI; Site.2: EcoRI;
First strand cDNA was primed with an anchored
XhoI-oligo(dT) primer [5'GGAGACTCGAGCGCGCGAGGAG(T)VN
3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand
was primed with a BamHI-dC primer
[5'AGAGACTCGATCGCGCGCGCAATAATAT(C) 3'].
Double-stranded cDNA was then digested with BamHI and XhoI
and directionally cloned into the BamHI and SalI sites of
lambda pSB vector. Library went through one round of
normalization. Library was constructed by Wei Yu at RIKEN
of Japan (Carninci, P., Westover, A., Nishiyama, Y., Ohsumi, T.,
Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M.,
Schneider, C., Hayashizaki, Y., High efficiency selection of
full-length cDNA by improved biotinylated cap trapper.,
DNA Res 4: 1, 61-6, Feb 28, 1997)"

ORIGIN
Query Match 27.9%; Score 370; DB 10; Length 497;
Best Local Similarity 100.0%; Pred. No. 2.6e-56;
Matches 370; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 589 CCAGCGCAGATTTCATCATCTCTGTGTGACCGAGAAAGTGCCTCCCATCTGCAGGAAGC 648
Db 128 CCAGCGCAGATTTCATCATCTCTGTGTGACCGAGAAAGTGCCTCCCATCTGCAGGAAGC 187
Qy 649 CACTATGCCAGAAAGCTGCTGACTGCGAGACTAGGCTCCCTCTGCGCCACGGTCCGTGCCAG 708
|||||

```

```

Db 188 CACTATGCCAGAAAGCTGCTGACTGCGAGAACTAGGCTCCCTCTGCGCACGGTCCGTGCCAG 247
Qy 709 CCAATAGATGCTCTGAGGCGCTGCCCTCTCCCACTTCACTCAGTTCACCAATCTAAATTT 768
|||||
Db 248 CCAATAGATGCTCTGAGGCGCTGCCCTCTCCCACTTCACTCAGTTCACCAATCTAAATTT 307
Qy 769 TTCAACAGAGATCTCTGTTTGGGGAACTTAAGTCAGATCCAGAACCTTTGGCTGCAAGGAG 828
|||||
Db 308 TTCAACAGAGATCTGTTTGGGGAACTTAAGTCAGATCCAGAACCTTTGGCTGCAAGGAG 367
Qy 829 TCTGGGAAATGTCAATTTCCCTAGAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTCG 888
|||||
Db 368 TCTGGGAAATGTCAATTTCCCTAGAGGAAGTTAGGGTGGGTGGAGCAAGCCCCACCTCG 427
Qy 889 TTTTTCGCCACAGCATCCCAATCGTGAAGAACTCGGAGAGGGTGGAGTCCACATCTAGG 948
|||||
Db 428 TTTTTCGCCACAGCATCCCAATCGTGAAGAACTCGGAGAGGGTGGAGTCCACATCTAGG 487
Qy 949 GTTGTCCTGCG 958
Db 488 GTTGTCCTGCG 497

RESULT 26
CA434086/c
LOCUS CA434086
DEFINITION UI-H-DF0-arq-p-13-0-UI.s2 NCI_CGAP_DF0 Homo sapiens cDNA clone
CA434086
ACCESSION CA434086.1 GI:24798506
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 424)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-x@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source
Location/Qualifiers
1..424
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-DF0-arq-p-13-0-UI"
/tissue_type="Subchondral Bone"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_DF0"
/notes="Organ: Bone; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site.1: EcoR I; Site.2: Not I;
NCI_CGAP_DF0 is a cDNA library containing the following
tissue(s): Subchondral Bone. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is GTTAAGCGTC.

```

FEATURES

High quality sequence stop: 301.

FEATURES
source
1. 304
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2064442"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/clone_lib="Soares NhMPu_S1"
/note="Organ: mixed (see below); Vector: pVT73D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not I;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NbHM, pregnant uterus
NbHPU, and fetal heart NbH19W) were mixed, and ss circles
were used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484498-489479."

ORIGIN

Query Match 22.9%; Score 304; DB 9; Length 304;
Best Local Similarity 100.0%; Pred. No. 1.6e-44;
Matches 304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 997 GAGTGGGCTGCAAGACTGAGCTAAATGTCCTCCCGGCTTGACTTTCTTTCTAGTCCT 1056
Db 304 GAGTGGGCTGCAAGACTGAGCTAAATGTCCTCCCGGCTTGACTTTCTTTCTAGTCCT 245
Qy 1057 GGGCCCTAGATTCGTGCACTTGGGTCTCTGACACACACACATCCCAAGTAGCCGAA 1116
Db 244 GGGCCCTAGATTCGTGCACTTGGGTCTCTGACACACACACATCCCAAGTAGCCGAA 185
Qy 1117 GAGCTAAACACAGGGGTTCCTAAATGGCTGCCCGCCACCCGGGCTCCCTTGGGCA 1176
Db 184 GAGCTAAACACAGGGGTTCCTAAATGGCTGCCCGCCACCCGGGCTCCCTTGGGCA 125
Qy 1177 AAAGGAATGTCAGCCCTACCCCAACCTTCACTACAGATCTGGGCGACCCAGAG 1236
Db 124 AAAGGAATGTCAGCCCTACCCCAACCTTCACTACAGATCTGGGCGACCCAGAG 65
Qy 1237 TATTTTATTTAAATGTCGCCATTTATGAGTATGATCAATTTGTATTAAATTAAG 1296
Db 64 TATTTTATTTAAATGTCGCCATTTATGAGTATGATCAATTTGTATTAAATTAAG 5
Qy 1297 TTAC 1300
Db 4 TTAC 1

RESULT 29
BE245533/c
LOCUS
DEFINITION
Baylor-HGSC project=TCBA Homo sapiens cDNA clone TCBAP3276, mRNA
sequence.

ACCESSION
BE245533
VERSION
1
KEYWORDS
EST.
SOURCE
Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE

1 (bases 1 to 376)

AUTHORS

Wei, Y., Tsang, Y.T.M., Mei, G., Xu, J.M., Ali-Osman Jr., F.R.,
Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.

TITLE

Pediatric Leukemia cDNA Sequencing Project

JOURNAL

Unpublished (2000)

COMMENT

Contact: Dr. Judith F. Margolin
Texas Children's Cancer Center and Human Genome Sequencing Center

at Baylor College of Medicine
1102 Bates, MC3-3320 Houston, TX 77030, USA
Tel: 832-824-4536

Fax: 832-825-4038

Email: clones@xccc.org

Citation: Carninci, P. and Hayashizaki, Y. High efficiency

full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)

Seq primer: M13 primer.

Location/Qualifiers

source
1. 376

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="TCBAP3276"

/sex="male"

/tissue_type="leukopheresis"

/cell_type="pre-B cell"

/dev_stage="pediatric 2 years"

/lab_host="DH10B"

/clone_lib="Pediatric pre-B cell acute lymphoblastic

leukemia Baylor-HGSC project=TCBA"

/note="Vector: lambda PSB; Site_1: BamHI; Site_2: EcoRI;

First strand cDNA was primed with an anchored

XhoI-oligo(dT) primer [5'GGAGACTCGAGCGCGGAGGAG(T)VN

3'; V=A,C,G; N=A,C,G,T] and then dG tailed. Second strand

was primed with a BamHI-dC primer

[5'AGAGCTCGATCCGCGCGGAATTAATAT(C) 3'].

Double-stranded cDNA was then digested with BamHI and XhoI

and directionally cloned into the BamHI and SalI sites of

lambda PSB vector. Library went through one round of

normalization. Library was constructed by Wei Yu at RIKEN

of Japan (Carninci, P., Westover, A., Nishiyama, Y., Ohsumi, T.,

Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M.,

Schneider, C., Hayashizaki, Y., High efficiency selection of

full-length cDNA by improved biotinylated cap trapper.,

DNA Res 4: 1, 61-6, Feb 28, 1997)"

ORIGIN

Query Match 22.5%; Score 299; DB 10; Length 376;
Best Local Similarity 100.0%; Pred. No. 9.6e-44;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 927 GAGGTGGAGTCCACATCTAGGTTGTCCTGCCCTTGGCTCTATCCCTGCCAGAGTG 986
Db 376 GAGGTGGAGTCCACATCTAGGTTGTCCTGCCCTTGGCTCTATCCCTGCCAGAGTG 317
Qy 987 GGAACCTGGAGAGTGGGCTCAAGACTGAGCCTAAATGTCTCCCGGCTTGACTTTCT 1046
Db 316 GGAACCTGGAGAGTGGGCTCAAGACTGAGCCTAAATGTCTCCCGGCTTGACTTTCT 257
Qy 1047 TTCTAGTCCTGGGCTAGATTCTGCACTTGGGCTCTCTGACACACACACATCCCAA 1106
Db 256 TTCTAGTCCTGGGCTAGATTCTGCACTTGGGCTCTCTGACACACACACATCCCAA 197
Qy 1107 GTAGCCGGAAGAGCTAAACACAGGGGTTCTTAAATGGCTGCCCCGCCACCCGGGCT 1166
Db 196 GTAGCCGGAAGAGCTAAACACAGGGGTTCTTAAATGGCTGCCCCGCCACCCGGGCT 137
Qy 1167 CCCTTGGGCAAAAGGAATTCTCAGCCCTACCCCAACCCCTTCAACTACCAATCTGGGC 1225
Db 136 CCCTTGGGCAAAAGGAATTCTCAGCCCTACCCCAACCCCTTCAACTACCAATCTGGGC 78

RESULT 30

AA644653

LOCUS

DEFINITION

af7411.r1 Soares NhMPu_S1 Homo sapiens cDNA clone IMAGE:1047789

5', mRNA sequence.

ACCESSION

AA644653

VERSION

AA644653.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

AA644653

LOCUS

DEFINITION

af7411.r1 Soares NhMPu_S1 Homo sapiens cDNA clone IMAGE:1047789

5', mRNA sequence.

ACCESSION

AA644653

VERSION

AA644653.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 385)

AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kubacka, I., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, E., Wylie, T., Waterston, R. and Wilson, R.

TITLE WashU-NCI human EST Project

JOURNAL Unpublished (1997)

COMMENT Contact: Wilton RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 499 Std Error: 0.00 Seq primer: -28ml3 rev2 ET from Amersham.

FEATURES source Location/Qualifiers 1. 385 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:1047789" /tissue types="Pooled human melanocyte, fetal heart, and pregnant uterus" /lab_host="DH10B" /clone_lib="Soares NHPu_S1" /notes="Organ: mixed (see below); Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2N8HM, pregnant uterus NHPU, and fetal heart N8H19W) were mixed, and as circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

ORIGIN Query Match 21.6%; Score 286; DB 9; Length 385; Best Local Similarity 100.0%; Pred. No. 1.8e-41; Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 967 TCTATCCCTGCCAGAGTGGGAACCTGGAGAGTGGGCTCCAGACTGAGCCCTAAATGTC 1026
DB 100 TCTATCCCTGCCAGAGTGGGAACCTGGAGAGTGGGCTCCAGACTGAGCCCTAAATGTC 159
QY 1027 TCCCCGGCCTTGACTTTCTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGCTCTCG 1086
DB 160 TCCCCGGCCTTGACTTTCTTCTAGTCCTGGGGCCTAGATTCTGCACTTGGGCTCTCG 219
QY 1087 ACACACACACACATCCCAAGTAGTACCGGAGAGCTAAACACAGGGGGTCTTAAATGGC 1146
DB 220 ACACACACACACATCCCAAGTAGTACCGGAGAGCTAAACACAGGGGGTCTTAAATGGC 279
QY 1147 TGGCCCCGCCACCGGGCTCCCTTGGGCGAAGAGTAATGTGACGCCCTACCCCAACCCCT 1206
DB 280 TGGCCCCGCCACCGGGCTCCCTTGGGCGAAGAGTAATGTGACGCCCTACCCCAACCCCT 339
QY 1207 CAACTACCAAGATCTGGGCCACCCAGCAGTATTTTATTATTAAT 1252
DB 340 CAACTACCAAGATCTGGGCCACCCAGCAGTATTTTATTATTAAT 385

RESULT 31
AW139822/c 434 bp mRNA linear EST 30-OCT-1999
LOCUS UI-H-B11-adw-c-09-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone
DEFINITION IMAGE:2718328 3', mRNA sequence.
ACCESSION AW139822

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AW139822.1 GI:6144540
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 434)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov

FEATURES source Location/Qualifiers 1. 434 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:2718328" /lab_host="NCI CGAP Sub3" /notes="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The NCI_CGAP_Sub3 library is a subtracted library derived from the NCI_CGAP_Sub1 library, which is a subtracted library derived from Bi. Bi constitutes a mixture of 21 normalized or subtracted NCI CGAP libraries: NCI CGAP Co4, NCI CGAP Pr22, NCI CGAP Pr28, NCI CGAP Co10, NCI CGAP Co16, NCI CGAP Kid5, NCI CGAP Kid12, NCI CGAP Kid3, NCI CGAP Kid11, NCI CGAP Lym2, NCI CGAP Br2, NCI CGAP Co8, NCI CGAP CLL1, NCI CGAP Lu24, NCI CGAP Lu19, NCI CGAP Co4, NCI CGAP Co6, NCI CGAP Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below: NCI CGAP Kid3 pool 1 LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE Clones 1322376-1323911, 1456008-1456775, 1500552-1502855); NCI CGAP Kid5 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clones 1323912-1325831, 1471368-1472903, 1492104-1493255); NCI CGAP Lu5 pool 1 LLAM 3575-3582, 3851-3854 (IMAGE Clones 1414920-1417991, 1520904-1522439); NCI CGAP Co4 pool 1 LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE Clones 1257096-1258631, 1469064-1470983, 1475592-1476743); NCI CGAP Pr22 pool 1 LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clones 905608-986759, 1101192-1101959, 1217928-1220615); NCI CGAP Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE Clones 1057416-1061255, 114584-1145351). Subtraction was performed as previously described [Bonaldo, Leunon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806. TAG TISSUE=brain TAG LIB=NCI CGAP_Brn23 TAG_SEQ=ATATC"]

ORIGIN Query Match 21.5%; Score 285; DB 10; Length 434; Best Local Similarity 99.7%; Pred. No. 2.4e-41; Matches 335; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 973 CTGTCCCGAGGTGGGAACCTGGAGTGGGCTGCAAGACTGAGCCTAATGTCCTCCCG 1032

```

|||
355 CCTGCCACAGGTGGGAACTGGAGAGTGGGCTGCAGACTGAGCCATAATGCTCCCGC 296
|||
1033 GCCTTGACTTTCTTTCTAGTCCTGGGCTTAGATTCTGCACTTGGGCTCTCTGACAAA 1092
|||
295 GCCTTGACTTTCTTTCTAGTCCTGGGCTTAGATTCTGCACTTGGGCTCTCTGACAAA 236
|||
1093 CACACCATCCCAAGTAGCCGAGAGCTAAACACAGGGGTTCTTAAATGGCTGCCCC 1152
|||
235 CACACCATCCCAAGTAGCTGGAAGAGCTAAACACAGGGGTTCTTAAATGGCTGCCCC 176
|||
1153 CGCCACCCGGGCTCCCTTGGGCAAAAGGAATGTCAGCCCTACCCCAACCCCTCAACTA 1212
|||
175 CGCCACCCGGGCTCCCTTGGGCAAAAGGAATGTCAGCCCTACCCCAACCCCTCAACTA 116
|||
1213 CCAGAACTTGGGCAACCCAGCAGTATTTATTTAAATGTTGCCCATTTATGAGTTA 1272
|||
115 CCAGAACTTGGGCAACCCAGCAGTATTTATTTAAATGTTGCCCATTTATGAGTTA 56
|||
1273 TGATCAATTTGTTATTAATTAAGCTTACAGATGCA 1308
|||
55 TGATCAATTTGTTATTAATTAAGTTACAGATGCA 20
|||

```

```

RESULT 32
H47130          431 bp      mRNA      linear      EST 16-AUG-1995
LOCUS          yp73a07.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
DEFINITION     IMAGE:193044 5', mRNA sequence.
ACCESSION     H47130
VERSION
KEYWORDS
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens

```

```

REFERENCE
AUTHORS
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.

```

```

TITLE
JOURNAL
COMMENT
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 935
High quality sequence stops: 322
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 935      Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 322.
Location/Qualifiers
1. .431
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3762065"
/db_xref="taxon:9606"
/clone="IMAGE:193044"
/sex="male"
/dev_stages="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGGAGAATTAATAAAGATCTTTTTTTTTTTTTTT 3']

```

FEATURES

```

source
1. .431
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3762065"
/db_xref="taxon:9606"
/clone="IMAGE:193044"
/sex="male"
/dev_stages="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
[5' AACTGGAGAATTAATAAAGATCTTTTTTTTTTTTTTT 3']

```

double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

```

Query Match      21.3%; Score 283; DB 14; Length 431;
Best Local Similarity 100.0%; Pred. No. 5.3e-41;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 466 GCAGAAACCAAGAGCCACTTGTAGCTGGTTTAAATTAGACAAGGATTTACTACCTGGCCCC 525
|||
Db 1 GCAGAAACCAAGAGCCACTTGTAGCTGGTTTAAATTAGACAAGGATTTACTACCTGGCCCC 60
|||
Qy 526 TGTGGCTTGCAGAAATTTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGGAATTTCCAGAA 585
|||
Db 61 TGTGGCTTGCAGAAATTTGTTGGAAGAGCTGGAGAAGCAGACTCTGCTGGAATTTCCAGAA 120
|||
Qy 586 CTCCAGCGCCAGATTTCATCATGTCCTTGTGACGAGAAAGCTGCCCATCTGCAGGA 645
|||
Db 121 CTCCAGCGCCAGATTTCATCATGTCCTTGTGACGAGAAAGCTGCCCATCTGCAGGA 180
|||
Qy 646 AGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCAGGTCGGTGC 705
|||
Db 181 AGCCACTATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCAGGTCGGTGC 240
|||
Qy 706 CAGCCAATAGATGTCCTGAGGCCCTGCCCTCTCCCACTTCACT 748
|||
Db 241 CAGCCAATAGATGTCCTGAGGCCCTGCCCTCTCCCACTTCACT 283
|||

```

RESULT 33

```

H78426
LOCUS          H78426
DEFINITION     Y422C04.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:234534 5', mRNA sequence.
ACCESSION     H78426
VERSION
KEYWORDS
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens

```

```

REFERENCE
AUTHORS
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.

```

TITLE

```

JOURNAL
COMMENT
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 929
High quality sequence stops: 356
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 929      Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 356.
Location/Qualifiers
1. .440
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3787711"
/db_xref="taxon:9606"
/clone="IMAGE:234534"

```

FEATURES

```

source
1. .440
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3787711"
/db_xref="taxon:9606"
/clone="IMAGE:234534"

```


Insert Length: 850 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 464.

FEATURES

source
1. .543
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2567906"
/lab_host="DH10B"
/clone_lib="Soares NFL T GBC S1"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NSHL19W, testis NHT, and B-cell NCI_CGAP_GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 20.9%; Score 277; DB 9; Length 543;
Best Local Similarity 99.1%; Pred. No. 4.8e-40;
Matches 527; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 772 CAAGAGATTCGTGGGGAACTTAAGTCAGATCCAGAACTTTGGCTGCAAGGAGTCT 831
Db 543 CAAGAGATTCGTGGGGAACTTAAGTCAGATCCAGAACTTTGGCTGCAAGGAGTCT 484
QY 832 GGGAAATGTCATTTCCCTAGAGGAAGTTAGGTGGGTGGAGAGAGCCCTCGGTTT 891
Db 483 GGGAAATGTCATTTCCCTAGAGGAAGTTAGGTGGGTGGAGAGAGCCCTCGGTTT 424
QY 892 TTCTGCCACAGCATCCATCTGAAGAACTCGGAGAGGGTGGAGTCCACATCTAGGGTT 951
Db 423 TTTTGGCCACAGCATCCATCTGAAGAACTCGGAGAGGGTGGAGTCCACATCTAGGGTT 364
QY 952 GTCTGCCCCCTGGCTTCATCCCTGCCAGAGGTGGAACTGGAGGATGGGCTGCAAGA 1011
Db 363 GTCTGCCCCCTGGCTTCATCCCTGCCAGAGGTGGAACTGGAGGATGGGCTGCAAGA 304
QY 1012 CTGAGCCTAAATGTCCTCCCGGCTTGATTTCTTCTAGTCTCTGGGCTAGATTCG 1071
Db 303 CTGAGCCTAAATGTCCTCCCGGCTTGATTTCTTCTAGTCTCTGGGCTAGATTCG 244
QY 1072 CACTTGGGGTCTCTGACACAAACACACCATCCCAAGTAGCCGGAAGAGCTAAACACAGGG 1131
Db 243 CACTTGGGGTCTTTGACACAAACACACCATCCCAAGTAGCCGGAAGAGCTAAACACAGGG 184
QY 1132 GGTTCTTAAATAGTGTCGCCCGCCACCCGGGCTCCCTTGGGCAAAAGAAATGTCAGC 1191
Db 183 GGTTTAAATAGTGTCGCCCGCCACCCGGGCTCCCTTGGGCAAAAGAAATGTCAGC 124
QY 1192 CCTACCCCAACCTTCACTACAGAACTCTGGGCCACCCGACAGTATTTTATTTAAA 1251
Db 123 CCTACCCCAACCTTCACTACAGAAATTTGGGCCACCCGACAGTATTTTATTTAAA 64
QY 1252 TGTGGCCCATTTATGAGTTATGATCAATTTGTTATTAATTAAGTTACAGA 1303
Db 63 TGTGGCCATTTATGAGTTATGATCAATTTGTTATTAATTAAGTTACAGA 12

RESULT 36
T85444
LOCUS
DEFINITION Yd82c03.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE:114724 5', mRNA sequence.
ACCESSION T85444
VERSION T85444.1 GI:713796

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 389)
AUTHORS Hallier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 948
High quality sequence stops: 261 Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 948 Std Error: 0.00
Seq primer: M13Rp1
High quality sequence stop: 261.
FEATURES
source
Location/Qualifiers
1. .389
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:470341"
/db_xref="taxon:9606"
/clone="IMAGE:114724"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site 1: Pac I; Site 2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5', AACTGGAAGAATTAATAAGATCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."

ORIGIN

Query Match 20.3%; Score 269; DB 14; Length 389;
Best Local Similarity 100.0%; Pred. No. 1.7e-38;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 467 CAGAAAAACAAGAGCCATTGTAGCTGTTTAATTAGACAAGATTACTACCTGGCCCT 526
Db 1 CAGAAAAACAAGAGCCATTGTAGCTGTTTAATTAGACAAGATTACTACCTGGCCCT 60
QY 527 GGTGGCTTGCAAAATCTTGAAGAGCTGGAGAGCAGACTCTGCTGAATTTCCAGGAAC 586
Db 61 GGTGGCTTGCAAAATCTTGAAGAGCTGGAGAGCAGACTCTGCTGAATTTCCAGGAAC 120
QY 587 TCCAGAGCCGAGATTTCATCTGTGTGTGACAGGAAAGCTGCCCCCATCTTCAGGAA 646
Db 121 TCCAGAGCCGAGATTTCATCTGTGTGTGACAGGAAAGCTGCCCCCATCTTCAGGAA 180
QY 647 GCACATATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCGGTCC 706
Db 181 GCACATATGCCAGAAAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCGGTCC 240
QY 707 AGCCAAATAGATGTCCTGAGGCTGCCCT 735
Db 241 AGCCAAATAGATGTCCTGAGGCTGCCCT 269

```

RESULT 37
BE219390/c
LOCUS
DEFINITION
hv57c03.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:3177508 3',
mRNA sequence.
ACCESSION
BE219390
VERSION
BE219390.1 GI:8906708
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 470)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
infoimage.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 446.
FEATURES
Location/Qualifiers
1..470
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3177508"
/tissue_type="carcinoid"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Lu24"
/notes="Organ: lung; Vector: p773D-Pac (Pharmacia) with a
modified polylinker. Plasmid DNA from the normalized
library NCI_CGAP Lu5 was prepared, and ss circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (clones
1414920-1417991 and 152904-1522439). Subtraction by Bento
Soares and M. Fatima Bonaldo. "
```

Query Match 20.3%; Score 269; DB 10; Length 470;

Best Local Similarity 100.0%; Pred. No. 1.4e-38; Mismatches 0; Indels 0; Gaps 0;

Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

988 GAATGGAGAGTGGGCTGCAAGACTGAGCTAAATGCTCCCGGGCTTGACTTTTCTT 1047

318 GAATGGAGAGTGGGCTGCAAGACTGAGCTAAATGCTCCCGGGCTTGACTTTTCTT 259

1048 TCTAGTCTGGGCTAGATTCTGCACTTGGGCTCTTGACACAAACACATCCCAAG 1107

258 TCTAGTCTGGGCTAGATTCTGCACTTGGGCTCTTGACACAAACACATCCCAAG 199

1108 TAGCCGGAAGAGCTAAACACAGGGGTTCTTAAATGGCTGCCCGCCACCGGGCTC 1167

198 TAGCCGGAAGAGCTAAACACAGGGGTTCTTAAATGGCTGCCCGCCACCGGGCTC 139

1168 CCTTGGCAAAAGAAATTGTCAGCCCTACCCCAACCTTCAACTACAGAAATCTGGGCCA 1227

138 CCTTGGCAAAAGAAATTGTCAGCCCTACCCCAACCTTCAACTACAGAAATCTGGGCCA 79

1228 CCCAGCAGTATTTTATTTAAATGTTG 1256

78 CCCAGCAGTATTTTATTTAAATGTTG 50

RESULT 38

BE245721/c

LOCUS

DEFINITION

TCBAP1D1925 Pediatric pre-B cell acute lymphoblastic leukemia

Baylor-HGSC project=TCBA Homo sapiens cDNA clone TCBAP1925, mRNA

sequence.

ACCESSION

BE245721

VERSION

BE245721.1 GI:9097468

SOURCE

EST.

ORGANISM

Homo sapiens (human)

REFERENCE

1 (bases 1 to 337)

Wei, Y., Tsang, Y.T.M., Mei, G., Ku, J.M., Ali-Osman Jr., F.R.,

Muzny, D., Bouck, J., Gibbs, R.A. and Margolin, J.F.

Pediatric Leukemia cDNA Sequencing Project

Unpublished (2000)

CONTACT: Dr. Judith F. Margolin

Texas Children's Cancer Center and Human Genome Sequencing Center

at Baylor College of Medicine

1102 Bates, MC3-3320 Houston, TX 77030, USA

Tel: 832-824-4536

Fax: 832-825-4038

Email: clones@ccc.org

Citation: Carninci, P. and Hayashizaki, Y. High efficiency

full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)

Seq primer: M13 primer.

FEATURES

Location/Qualifiers

1..337

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="TCBAP1925"

/sex="male"

/tissue_type="leukopheresis"

/cell_type="pre-B cell"

/dev_stage="pediatric 2 years"

/lab_host="DH10B"

/clone_lib="Pediatric pre-B cell acute lymphoblastic

leukemia Baylor-HGSC project=TCBA"

/notes="Vector: lambda pSB; Site 1: BamHI; Site 2: EcoRI;

First strand cDNA was primed with an anchored

XhoI-oligo(dT) primer [5'-GGAGACTCGAGCGCCGAGAGAG(T)VN

3'; 5'-A,C,G; N=A,C,G,T] and then dG tailed. Second strand

was primed with a BamHI-dC primer

[5'-AGAGAGTCGATCCGCGCCGCAATAATAAT(C) 3'].

Double-stranded cDNA was then digested with BamHI and XhoI

and directionally cloned into the BamHI and SalI sites of

lambda pSB vector. Library went through one round of

normalization. Library was constructed by Wei Yu at RIKEN

of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,

Itoh M, Nagaoka S, Sasaki N, Okazaki Y, Muramatsu M,

Schneider C, Hayashizaki Y. High efficiency selection of

full-length cDNA by improved biotinylated cap trapper.,

ORIGIN

Query Match 19.6%; Score 260; DB 10; Length 337;

Best Local Similarity 100.0%; Pred. No. 7.3e-37;

Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

967 TCTATCCCTGCCAGAGTGGGAACCTGGAGAGTGGGCTGCAAGACTGAGCTTAATGTC 1026

336 TCTATCCCTGCCAGAGTGGGAACCTGGAGAGTGGGCTGCAAGACTGAGCTTAATGTC 277

1027 TCCCGGGCTTGACTTTTCTTCTAGTCTGGGCTAGATTCTGACTTTGGGCTCTCG 1086

276 TCCCGGGCTTGACTTTTCTTCTAGTCTGGGCTAGATTCTGACTTTGGGCTCTCG 217

1087 ACACAACACACCATCCCAAGTAGCGGAGAGCTAAACACACAGGGGTTCTTAAATGGC 1146

Db 216 ACACACACACATCCCAAAGTAGCGGAGAGAGCTAAACACAGCGGGGTCTCTTAAATAGC 157
 Qy 1147 TGCCCCCGCACCGGGCTCCCTTGGCCAAAGGAATTCAGCCCTTACCCCAACCTT 1206
 Db 156 TGCCCCCGCACCGGGCTCCCTTGGCCAAAGGAATTCAGCCCTTACCCCAACCTT 97
 Qy 1207 CAATACCAAGATCTGGGCC 1226
 Db 96 CAATACCAAGATCTGGGCC 77

RESULT 39
 LOCUS BF696473 749 bp mRNA linear EST 22-DEC-2000
 DEFINITION 602125145F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4282091 5', mRNA sequence.
 ACCESSION BF696473
 VERSION BF696473.1 GI:11981881
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 749)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: L1CM1111 row: j column: 12
 High quality sequence stop: 560.

FEATURES

Location/Qualifiers
 1..749

source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4282091"
 /tissue_type="primitive neuroectoderm"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC_56"
 /note="Organ: brain; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgcctcgcc); Site 2: SfiI (ggccattatggcc); SfiI (ggcgcctcgcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATTCTAGGCGGAGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 19.2%; Score 255; DB 10; Length 749;
 Best Local Similarity 100.0%; Pred. No. 2.5e-36;
 Matches 255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 70 ACCAGAACTATATGGTTCCTACTCTCCCACTGATCTGCTGGTCAGTGATGCTGTG 129
 Db 145 ACCAGAACTATATGGTTCCTACTCTCCCACTGATCTGCTGGTCAGTGATGCTGTG 204
 Qy 130 GCCTGTGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCACCACTTCC 189
 Db 205 GCCTGTGAAGGCACCTGGTAGTTGAGTCCACACATTATAGTCATGTGCCACCACTTCC 264
 Qy 190 TGCCACAGCGCGGAGGACAGGGGTAGGGGTATACCCAAAGCTGATGCAGAGCCCATTAGC 249

Db 265 TGCCCCAGCGCGGAGGACAGGGTAGGGGTATACCCAAAGCTGATGCAGAGCCCATTAGC 324
 Qy 250 CTAAGAACAATCGAGGACAAAGCCTCCCTGGATGATCGAGTCCCGAGTAGCTCTGAACA 309
 Db 325 CTAAGAACAATCGAGGACAAAGCCTCCCTGGATGATCGAGTCCCGAGTAGCTCTGAACA 384
 Qy 310 AGAGTCCAGCAACC 324
 Db 385 AGAGTCCAGCAACC 399

RESULT 40
 LOCUS BU754096 270 bp mRNA linear EST 10-OCT-2002
 DEFINITION UI-1-BB1P-atn-h-07-0-UI.s1 NCI_CGAP_P16 Homo sapiens cDNA clone UI-1-BB1P-atn-h-07-0-UI 3', mRNA sequence.
 ACCESSION BU754096
 VERSION BU754096.1 GI:23712667
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 270)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Steven Brown
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

Location/Qualifiers
 1..270

source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-1-BB1P-atn-h-07-0-UI"
 /tissue_type="Placenta"
 /dev_stage="Full Term"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI_CGAP_P16"
 /note="Organ: Placenta; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI_CGAP_P16 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tags for this library are GA, AGGAA. For additional information, contact: Bento Soares, bento-soares@uiowa.edu
 TAG_TISSUE=placenta human full term
 TAG_LIB=UI-1-BB1P
 TAG_SEQ=AGGAA"

ORIGIN

Query Match 18.6%; Score 247; DB 13; Length 270;
 Best Local Similarity 100.0%; Pred. No. 1.7e-34;
 Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1057 GGGCCCTAGATTCTGCACTTGGGTCTCTGACACACACACATCCCAAGTAGCCGGAA 1116

Db	262	GGGCGCTAGATTTCGTCACTTTGGGCTCTCTGCACACAACACCACCATCCCAAAGTAGCCGAA	203
Qy	1117	GAGCTAAACACAGGGGGTTCCTTAATAATGGCTGCCCGCCGACCGGGCCTCCCTTTGGGA	1176
Db	202	GAGCTAAACACAGGGGGTTCCTTAATAATGGCTGCCCGCCGACCGGGCCTCCCTTTGGGA	143
Qy	1177	AAAGGAATTGTCAGCCCTACCCCACCCCTTCAACTACAGAATCTGGSCCACCCACGACG	1236
Db	142	AAAGGAATTGTCAGCCCTACCCCACCCCTTCAACTACAGAATCTGGSCCACCCACGACG	83
Qy	1237	TATTTTATTTAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAATTTAAG	1296
Db	82	TATTTTATTTAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTATTAATTTAAG	23
Qy	1297	TTACAGA 1303	
Db	22	TTACAGA 16	
RESULT 41			
AW964518			
LOCUS	AW964518	378 bp mRNA linear EST 01-JUN-2000	
DEFINITION	EST376591 MAGE resequences, MAGH Homo sapiens cDNA, mRNA sequence.		
ACCESSION	AW964518		
VERSION	AW964518.1 GI:8154354		
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 378) Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C., Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and Quackenbush,J.		
TITLE	Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: John Quackenbush The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 3528 Fax: 301 838 0208 Email: johnd@tigr.org Plate: 202 Seq primer: Reverse. Location/Qualifiers 1..378 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone_lib="MAGE resequences, MAGH" /vector="Vector: pBluescriptSKm"		
FEATURES			
source			
ORIGIN			
Query Match	18.2%; Score 241; DB 10; Length 378;		
Best Local Similarity	100.0%; Pred.No.1.4e-33;		
Matches	241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	640	GCAGGAGCCACTATGCCAGAACCTCTGACTGCAGAACTAGCTCCCTCTGCGCAGGT	699
Db	1	GCAGGAGCCACTATGCCAGAACCTCTGACTGCAGAACTAGCTCCCTCTGCGCAGGT	60
Qy	700	CCGTGCCAGCAATAGATGTCCTGAGCGCTGCCCTTCCCACTTCACTCAGTTCCCAA	759
Db	61	CCGTGCCAGCAATAGATGTCCTGAGCGCTGCCCTTCCCACTTCACTCAGTTCCCAA	120
Qy	760	TCTAAATTTTACAAGAGATTCTGTTGGGGGAACCTTAAGTCAGATCCAGAACTTGCT	819
Db	121	TCTAAATTTTACAAGAGATTCTGTTGGGGGAACCTTAAGTCAGATCCAGAACTTGCT	180
Qy	820	GCAAGGAGTCTGGGAATTCATTTCCCTTAGAAGGAAGTAGGGTGGGTGGAGCAAGCC	879

QY 1271 TATGATCAATTGTTAAATTAAGTTACAGATGTC 1307
 |||||
 Db 37 TATGATCAATTGTTAAATTAAGTTACAGATGTC 1

RESULT 43

LOCUS BB817016 334 bp mRNA linear EST 21-SEP-2000
 DEFINITION RC0-BN0248-310700-025-f05 BN0248 Homo sapiens cDNA, mRNA sequence.
 BB817016
 VERSION BB817016.1 GI:10249250
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 334)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 PUBMED 10737800

COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 103, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gechtml2.pl?cl=st2-RC0-BN0248-310700-025-f05&t3=2000-07-31&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 26
 High quality sequence stop: 334.
 Location/Qualifiers

FEATURES
 source
 1..334
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_libs="BN0248"
 /note="Organ: breast_normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORFESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 17.0%; Score 225; DB 10; Length 334;
 Best Local Similarity 100.0%; Pred. No. 9.9e-31;
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 83 TGGTTCCTACTTCTCCCACTGATCTGCTGGTCAAGTATGATGCTGGCCTGTGGAAGGC 263
 Db 322 TGGTTCCTACTTCTCCCACTGATCTGCTGGTCAAGTATGATGCTGGCCTGTGGAAGGC 263
 QY 143 ACCTGGTATGAGTCCACATATATAGTATGTCACACACCTCTCTCCGCCACAGGCCG 202
 Db 262 ACCTGGTATGAGTCCACATATATAGTATGTCACACACCTCTCTCCGCCACAGGCCG 203
 QY 203 AGGCAGGGTGAGGTATACCCAAAGCTGATGCAGAGCCCATAGCCCTAAAGCAACTG 262

Db 202 AGGCAGGGTGAGGTATACCCAAAGCTGATGCAGAGCCCATAGCCTAAAGCAACTG 143
 QY 263 CAGGACAAGCCTCCCTGGATGATCGAGGTCCCAAGTAGCTCTGAA 307
 |||||
 Db 142 CAGGACAAGCCTCCCTGGATGATCGAGGTCCCAAGTAGCTCTGAA 98

RESULT 44

LOCUS BG912772 405 bp mRNA linear EST 05-JUN-2001
 DEFINITION 602807993F1 NCI CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4940254
 5', mRNA sequence.

ACCESSION BG912772
 VERSION BG912772.1 GI:14293248
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 405)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: David N. Louis, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LUAM10878 row: i column: 23
 High quality sequence stop: 405.
 Location/Qualifiers

FEATURES

source
 1..405
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4940254"
 /tissue_type="anaplastic oligodendroglioma with lp/19q loss"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NCI CGAP_Brn67"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.3 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library."

ORIGIN

Query Match 16.7%; Score 221; DB 12; Length 405;
 Best Local Similarity 99.6%; Pred. No. 4.1e-30;
 Matches 271; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1037 TGACTTTCTTTCTAGTCTCGGGCTAGATTCTGCACCTGGGGTCTCTGCACACACACA 1096
 Db 116 TGACTTTCTTTCTAGTCTCGGGCTAGATTCTGCACCTGGGGTCTCTGCACACACA 175
 QY 1097 CCATCCCAAGTAGCCGGAAGAGCTAAACACAGGGGTTCTTAAATGGTCCCGCCG 1156
 Db 176 CCATCCCAAGTAGCCGGAAGAGCTAAACACAGGGGTTCTTAAATGGTCCCGCCG 235
 QY 1157 ACCGGGCTCCCTGGGCAAAAGGAATGTCAGCCCTACCCCAACCTTCACTACACAG 1216
 Db 236 ACCGGGCTCCCTGGGCAAAAGGAATGTCAGCCCTACCCCAACCTTCACTACACAG 295
 QY 1217 AATCTGGGCCACCCAGCAGTATTTTATTTAAATGTGCCCATTTTATGAGTTATGAT 1276
 Db 296 AATCTGGGCCACCCAGCAGTATTTTATTTAAATGTGCCCATTTTATGAGTTATGAT 355
 QY 1277 CAATTGTATTAATTAAGTTACAGATGTC 1308
 |||||


```
Db 356 CAATTGTGTTAAATTAAGTTACAGATGTCA 387

RESULT 45
AL554421
LOCUS AL554421 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA EST 31-MAY-2003
DEFINITION clone CS0D1084YE24 5-PRIME, mRNA sequence.
ACCESSION AL554421
VERSION AL554421.2 GI:31276232
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 768)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:12895186.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 221.r For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0D1084BC12Qp1&cluster=221.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0D1084BC12Qp1.
FEATURES
source
1. 768
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1084YE24"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 16.78; Score 221; DB 9; Length 768;
Best Local Similarity 98.94; Pred. No. 2.2e-30;
Matches 471; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 831 TGGGAATGTCATTTCCTAGAGGAAGTTAGGTGGTGGAGCGAGCCACCTCGGTT 890
Db 269 TGGGAATGTCATTTCCTAGAGGAAGTTAGGTGGTGGAGCGAGCCACCTCGGTT 328
QY 891 TTCTGCGCAGCATCCATCGTGAAGAACTCGGGAGAGGTGGAGTCCACATCTAGGTT 950
Db 329 TTCTGCGCAGCATCCATCGTGAAGAACTCGGGAGAGGTGGAGTCCACATCTAGGTT 388
QY 951 TGTCCCTGCCCTTGGCTCTATCCCTGCCAGAGTGGGAACCTGGAGAGTGGGCTGCAAG 1010
Db 389 TGTCCCTGCCCTTGGCTCTATCCCTGCCAGAGTGGGAACCTGGAGAGTGGGCTGCAAG 448
QY 1011 ACTGAGCGCTAAATGTCTCCCGGCGCTTGAATTTCTTCTAGTCCTGGGCGCTAGATTCT 1070
Db 449 AMTGAGCCTAAATGTCTCCCGGCGCTTGAATTTCTTCTAGTCCTGGGCGCTAGATTCT 508
QY 1071 GCATTTGGGTCTCTGACACACACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGG 1130
Db 509 GCATTTGGGTCTCTGACACACACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGG 568
QY 1131 GGGTTCTTAAATGGCTGCCCGCCGACACCGGGCGCTCCCTTTGGGCAAAAGGAATGTGAC 1190
Db 569 GGGTTCTTAAATGGCTGCCCGCCGACACCGGGCGCTCCCTTTGGGCAAAAGGAATGTGAC 628

QY 1191 CCTTACCCCAACCTTCACTACAGATCTGGGCCACCCAGAGTATTTTATTTAA 1250
Db 629 CACTACCCCAACCTTCACTACAGATCTGGGCCACCCAGAGTATTTTATTTAA 688
QY 1251 ATGTTGCCATTTTATGAGTTATGATCAATTTGTTATTAAATTAAGTTACAGATGT 1306
Db 689 ATGTTGCCATTTTATGAGTTATGATCAATTTGTTATTAAATTAAGTTACAGATGT 744

RESULT 46
BU527114
LOCUS BU527114 931 bp mRNA linear EST 13-SEP-2002
DEFINITION AGENCOURT 10155963 NIH_MGC_101 Homo sapiens cDNA clone
IMAGE:6536868 5', mRNA sequence.
ACCESSION BU527114
VERSION BU527114.1 GI:22837555
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 931)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2698 row: 0 column: 12
High quality sequence stop: 552.
FEATURES
source
1. 931
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6536868"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC 101"
/note="Organ: lung; Vector: pOTB7; Site 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
ORIGIN
Query Match 16.0%; Score 212; DB 13; Length 931;
Best Local Similarity 100.0%; Pred. No. 6.9e-29;
Matches 212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACAGTGGTCACAGGTAGTACCTGGTCTAGGTTGGCTGAGAGCCACCTCTCTGC 60
Db 401 TCACAGTGGTCACAGGTAGTACCTGGTCTAGGTTGGCTGAGAGCCACCTCTCTGC 460
QY 61 CACCCCCACACCAAGAACTATATGGTTCTCTCTCCCACTGATCTGCTGTCAGTGAT 120
Db 461 CACCCCCACACCAAGAACTATATGGTTCTCTCTCCCACTGATCTGCTGTCAGTGAT 520
QY 121 GATGCTGTGGCTGTGGAAGGACCTGTGTAGTTGAGTCCACACATTATAGTCATGTGCCA 180
Db 521 GATGCTGTGGCTGTGGAAGGACCTGTGTAGTTGAGTCCACACATTATAGTCATGTGCCA 580
QY 181 CCACCTTCTCTGCCACAGGCCGAGGACAGG 212
```

```

Db      581 CCACCTTCTCCACAGGCGGACAGG 612
|||||
RESULT 47
AI864353/c
LOCUS      310 bp      mRNA      linear      EST 30-AUG-1999
DEFINITION w978f06.xl Soares NSF F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
IMAGE:2371235 3', mRNA sequence.
ACCESSION  AI864353
VERSION     AI864353.1  GI:5528460
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 310)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs@mail.nih.gov
          This clone is available royalty-free through LNL; contact the
          IMAGE Consortium (info@image.llnl.gov) for further information.
          Seq primer: -40UP from Gibco
          High quality sequence stop: 301.
FEATURES   Location/Qualifiers
            1..310
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:2371235"
             /lab_host="DH10B"
             /clone_lib="Soares NSF F8_9W_OT_PA_P_S1"
             /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
             a modified polylinker; Site 1: Not I; Site 2: Eco RI;
             Equal amounts of plasmid DNA from five normalized
             libraries were mixed, and ss circles were made in vitro.
             Following HAP purification, this DNA was used as tracer in
             a subtractive hybridization reaction. The driver was
             PCR-amplified cDNAs from pools of 5,000 clones made from
             the same 5 libraries. The pools consisted of the following
             libraries and cloneIDs: Soares NbHSF pool 1:
             309384-310919, 323208-325895 Soares Nb2HP pool 1:
             145032-147335, 147720-148103, 148872-149255, 15002 -
             150407, 151176-152327 Soares Nb2HF8-9W pool 1:
             758280-760583, 772104-774407 Soares NbHPA pool 1:
             304776-306311, 320136-322823, 326280-326663 Soares NbHOR
             pool 1: 723720-726407, 739080-740999 Subtraction by Bento
             Soares and M. Fatima Bonaldo."
ORIGIN
Query Match      15.9%; Score 211; DB 9; Length 310;
Best Local Similarity 100.0%; Pred. No. 3e-28;
Matches 211; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1093 CACACATCCCAAGTAGCCGGAAGAGCTAAACACAGGGGTTCTTAAATGGCTGCC 1152
Db      212 CACACATCCCAAGTAGCCGGAAGAGCTAAACACAGGGGTTCTTAAATGGCTGCC 153
QY      1153 CGCCACCGGGCTCCCTTGGGCAAGAAATGTCAGCCCTACCCCAACCCCTTCAACTA 1212
Db      152 CGCCACCGGGCTCCCTTGGGCAAGAAATGTCAGCCCTACCCCAACCCCTTCAACTA 93
QY      1213 CCAGAAATCTGGGCCACCCACAGCAGTATTTTATTTAAATGTGGCCATTTATGAGTTA 1272
Db      92 CCAGAAATCTGGGCCACCCACAGCAGTATTTTATTTAAATGTGGCCATTTATGAGTTA 33
QY      1273 TGATCAATTTGATTTAAATTAAGTTACAGA 1303
Db      32 TGATCAATTTGATTTAAATTAAGTTACAGA 2

```

```

RESULT 48
BQ025918/c
LOCUS      270 bp      mRNA      linear      EST 27-MAR-2002
DEFINITION UI-1-BB1p-ayg-d-07-0-UI.s1 NCI-CGAP_P16 Homo sapiens cDNA clone
UI-1-BB1p-ayg-d-07-0-UI 3', mRNA sequence.
ACCESSION  BQ025918
VERSION     BQ025918.1  GI:19761197
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 270)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
          Email: cgapbs@mail.nih.gov
          Tissue Procurement: Dr. Steven Brown
          cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
          cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
          Clone Distribution: Clone distribution information can be obtained
          from Dr. M. Bento Soares, bento-soares@uiowa.edu
          Seq primer: M13 FORWARD
          POLYA-Yes.
FEATURES   Location/Qualifiers
            1..270
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="UI-1-BB1p-ayg-d-07-0-UI"
             /tissue_type="Placenta"
             /dev_stage="Full term"
             /lab_host="DH10B (Life Technologies)"
             /clone_lib="NCI CGAP_P16"
             /note="Organ: Placenta; Vector: pT7T3-Pac (Pharmacia) with
             a modified polylinker; Site 1: EcoR I; Site 2: Not I;
             NCI-CGAP_P16 is a subtracted cDNA library constructed
             according to Bonaldo, Lennon and Soares, Genome Research,
             6:791-806, 1996. First strand cDNA synthesis was primed
             with an oligo-dT primer containing a Not I site. Double
             stranded cDNA was ligated to an EcoR I adaptor, digested
             with Not I, and cloned directionally into pT7T3-Pac
             vector. The oligonucleotide used to prime the synthesis of
             first-strand cDNA contains a library tag sequence that is
             located between the Not I site and the (dr)18 tail. The
             sequence tags for this library are GA, AGGAA. For
             additional information, contact: Bento Soares,
             bento-soares@uiowa.edu
             TAG_TISSUE=placenta human full term
             TAG_LIB=UI-1-BB1p
             TAG_SEQ=AGGAA"
ORIGIN
Query Match      14.8%; Score 196; DB 12; Length 270;
Best Local Similarity 99.6%; Pred. No. 1.5e-25;
Matches 246; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1057 GGGGCTAGATTCTGCACCTTGGGGTCTCTGACACACACACACATCCCAAGTAGCGGAA 1116
Db      262 GGGGCTAGATTCTGCACCTTGGGGTCTCTGACACACACACACATCCCAAGTAGCGGAA 203
QY      1117 GAGCTAAACACAGGGGTTCTTAAATGGCTGCCCGCCCGGCTCCCTTGGGCA 1176
Db      202 GAGCTAAACACAGGGGTTCTTAAATGGCTGCCCGCCCGGCTCCCTTGGGCA 143
QY      1177 AAAGAAATGTCAGCCCTACCCCAACCCCTTCAATACCAAGAATCTGGGCCACCCACGAC 1236
Db      142 AAAGAAATGTCAGCCCTACCCCAACCCCTTCAATACCAAGAATCTGGGCCACCCACGAC 83

```

QY 1237 TATTTTATTTAAATGTTGCCCAATTTATGAGTTATGATCAATTTGTTATTAATTAAG 1296
 |||||
 Db 82 TATTTTATTTAAATGTTGCCCAATTTATGAGTTATGATCAATTTGTTATTAATTAAG 23
 |||||
 QY 1297 TTACAGA 1303
 |||||
 Db 22 TTACAGA 16
 |||||

RESULT 49
 AA037877/c
 LOCUS zf04f10.sl Soares fetal heart NbHH19W Homo sapiens linear EST 01-FEB-1997
 DEFINITION IMAGE:375979 3', mRNA sequence.
 ACCESSION AA037877
 VERSION
 KEYWORDS EST.
 SOURCE AA037877.1 GI:1513013
 ORGANISM Homo sapiens (human)

REFERENCE
 AUTHORS
 Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
 Chisoe,S., Dietrich,N., PuBuque,T., Favello,A., Gish,W.,
 Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
 Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
 Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J.,
 Trevasaki,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R.
 and Marra,M.

TITLE Generation and analysis of 280,000 human expressed sequence tags
 JOURNAL Genome Res. 6 (9), 807-828 (1996)
 MEDLINE 97044478
 PUBMED 8889549

COMMENT
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

This clone is available royalty-free through LINL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 Insert Length: 768 Std Error: 0.00
 Seq primer: -40M13 fwd. from Amersham
 High quality sequence stop: 376.

FEATURES
 Location/Qualifiers
 1..488
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:1284235"
 /db_xref="taxon:9606"
 /clone="IMAGE:375979"
 /sex="unknown"
 /dev_stage="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares fetal heart NbHH19W"

/notes="Organ: heart; Vector: p773D (Pharmacia) with a
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'
 TGTTACCAATCTGAAGTGGAGCGCGGCACATTTTTTTTTTTTTTTT 3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified p773 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by
 M.Fatma Bonaldo. This library was constructed from the
 same fetus as the fetal lung library, Soares fetal lung
 NbHH19W."

ORIGIN

Query Match 14.6%; Score 194; DB 9; Length 488;
 Best Local Similarity 99.6%; Pred. No. 1.9e-25;
 Matches 244; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1063 TAGATTCTGCATCTGGGTCTCTGACACACACACCATCCCAAGTAGCCGAGAGCTA 1122
 |||||
 Db 268 TAGATTCTGCATCTGGGTCTCTGACACACACACCATCCCAAGTAGCCGAGAGCTA 209
 |||||
 QY 1123 AACACAGGGGTTCTTAAATGGCTGCCCGCCGACACCGGGGCTCCCTTGGGCAAAAGGA 1182
 |||||
 Db 208 AACACAGGGGTTTAAATGGCTGCCCGCCGACACCGGGGCTCCCTTGGGCAAAAGGA 149
 |||||
 QY 1183 ATTGTAGCCCTACCCCAACCCCTTCAACTACAGAACTCTGGGCCACCCAGCAGTATTTT 1242
 |||||
 Db 148 ATTGTAGCCCTACCCCAACCCCTTCAACTACAGAACTCTGGGCCACCCAGCAGTATTTT 89
 |||||
 QY 1243 TATTTAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTTATTAATTAAGTTACAG 1302
 |||||
 Db 88 TATTTAAATGTTGCCCATTTTATGAGTTATGATCAATTTGTTATTAATTAAGTTACAG 29
 |||||
 QY 1303 ATGTC 1307
 |||||
 Db 28 ATGTC 24
 |||||

RESULT 50
 BM678923/c

LOCUS BM678923 401 bp mRNA linear EST 27-FEB-2002
 DEFINITION UI-E-E00-ahx-d-01-0-UI.sl UI-E-E00 Homo sapiens cDNA clone
 UI-E-E00-ahx-d-01-0-UI 3', mRNA sequence.

ACCESSION BM678923
 VERSION BM678923.1 GI:18988819
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS
 1 (bases 1 to 401)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548

COMMENT
 Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Forward
 POLYA=Yes.

FEATURES
 source

Location/Qualifiers
 1..401
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-E00-ahx-d-01-0-UI"
 /tissue_type="fetal eye"
 /dev_stage="fetal"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-E-E00"
 /notes="Organ: eye; Vector: p773-Pac (Pharmacia) with a
 modified polylinker; Site_1: EcoR I; Site_2: Not I;
 UI-E-E00 is a cDNA library containing the following
 tissue(s): fetal eye. The library was constructed
 according to Bonaldo, Lennon and Soares, Genome Research,
 6:791-806, 1996. First strand cDNA synthesis was primed
 with an oligo-dT primer containing a Not I site. Double
 stranded cDNA was ligated to an EcoR I adaptor, digested

with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tag for this library is GCGGTATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).
TAG_TISSUE=human fetal eye
TAG_LIB=UI-E-E00
TAG_SEQ=GCGGTATACC"

ORIGIN

Query Match		13.7%;	Score 182;	DB 12;	Length 401;
Best Local Similarity		99.1%;	Pred. No. 2.8e-23;		
Matches 332;		Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
Qy	969	TATCCCTGCCAGAGGTGGGAACTGGAGAGTGGCTGCAAGACTGAGCCTAAATGTCTC	1028		
Db	350	TATCCCTGCCAGAGGTGGGAACTGGAGAGTGGCTGCAAGACTGAGCCTAAATGTCTC	291		
Qy	1029	CCCGGCTTGACTTTTCTTCTTAGTCCTGGGGCTAGATTCTGCACCTGGGGTCTCTGAC	1088		
Db	290	CCCGGCTTGACTTTTCTTCTTAGTCCTGGGGCTAGATTCTGCACCTGGGGTCTCTGAC	231		
Qy	1089	ACAACACACCATCCCAAGTAGCCGGAAGAGCTAAACACAGGGGGTCTTAAATGGCTG	1148		
Db	230	ACAACACACCATCCCAAGTAGCCGGAAGAGCTAAACACAGGGGGTCTTAAATGGCTG	171		
Qy	1149	CCCCGCCACCGGGCCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCCTCA	1208		
Db	170	CCCCGCCACCGGGCCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCAACCCCTCA	111		
Qy	1209	ACTACAGAACTGGGGCCACCCAGCAGTATTTTATTATAAATGTGCCCAATTTATGA	1268		
Db	110	AATACCAAGATCTGGGCCACCCAGCAGTATTTTATTATAAATGTGCCCAATTTATGA	51		
Qy	1269	GTTATGATCAATTGTATTAAATTAAAGTTACAGA	1303		
Db	50	GTTATGATCAATTGTATTAAATTAAAGTTACAGA	16		

Search completed: September 17, 2004, 05:40:47
Job time : 3287 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 00:14:08 ; Search time 5150 Seconds
(without alignments)
11168.195 Million cell updates/sec

Title: US-09-927-091-3_COPY_2500_3826
Perfect score: 1327
Sequence: 1 tccacagtgcacaggtag.....aaaaaaaaaaaaaaaaaa 1327

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_bt.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_hg_hum.*
- 31: em_hg_inv.*
- 32: em_hg_other.*
- 33: em_hg_mus.*
- 34: em_hg_pln.*
- 35: em_hg_rtd.*
- 36: em_hg_nam.*
- 37: em_hg_vrt.*
- 38: em_sy.*
- 39: em_hgo_hum.*
- 40: em_hgo_mus.*
- 41: em_hgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
	1	1327	100.0	3259	9	BC001222	BC001222	Homo sapi
	2	1327	100.0	3259	9	BC007999	BC007999	Homo sapi
	3	1327	100.0	3259	9	BC011689	BC011689	Homo sapi
	4	1327	100.0	3259	9	BC012152	BC012152	Homo sapi
	5	1327	100.0	3826	6	AX491279	AX491279	Sequence
	6	1308	98.6	23433	6	AX491283	AX491283	Sequence
	7	1308	98.6	164950	2	AC022262	AC022262	Homo sapi
	8	1307	98.5	3243	6	AX775831	AX775831	Sequence
	9	1257	94.7	30676	6	AX491284	AX491284	Sequence
	10	1257	94.7	64693	9	AL662907	AL662907	Human DNA
	11	1256	94.6	3436	6	AX877131	AX877131	Sequence
	12	1256	94.6	3436	6	BD156501	BD156501	Primer fo
	13	1256	94.6	3436	9	AK001621	AK001621	Homo sapi
	14	1053	79.4	136095	2	AF161326	AF161326	Homo sapi
	15	726	54.7	30625	6	AX491281	AX491281	Sequence
	16	452	34.1	2207	6	AX775827	AX775827	Sequence
	17	452	34.1	2207	6	AX776017	AX776017	Sequence
	18	452	34.1	2246	6	AX775829	AX775829	Sequence
	19	423	31.9	557	6	AX873748	AX873748	Sequence
	20	423	31.9	557	6	BD153810	BD153810	Primer fo
	21	336	25.3	3148	9	AK122896	AK122896	Homo sapi
	22	62	4.7	250	11	GI5084	GI5084	human STS S
	23	29	2.2	614	6	AX388847	AX388847	Sequence
	24	29	2.2	1652	6	AX771586	AX771586	Sequence
	25	29	2.2	162959	2	AC019255	AC019255	Homo sapi
	26	29	2.2	178217	9	AC022080	AC022080	Homo sapi
	27	29	2.2	180338	10	AC090496	AC090496	Mus muscu
	28	28	2.1	723	10	BC060953	BC060953	Mus muscu
	29	28	2.1	136037	9	AC004104	AC004104	Homo sapi
	30	28	2.1	137648	2	AC024316	AC024316	Homo sapi
	31	28	2.1	158608	9	AC051635	AC051635	Homo sapi
	32	28	2.1	187557	2	AC069032	AC069032	Homo sapi
	33	28	2.1	193978	2	AC068678	AC068678	Homo sapi
	34	28	2.1	209462	2	AC011814	AC011814	Homo sapi
	35	27	2.0	4150	10	BC060625	BC060625	Mus muscu
	36	27	2.0	4685	9	HSN803417	HSN803417	Homo sapi
	37	27	2.0	88643	8	AC022522	AC022522	Arabidops
	38	27	2.0	105199	9	AP001576	AP001576	Homo sapi
	39	27	2.0	133897	9	AC112138	AC112138	Homo sapi
	40	27	2.0	152772	9	AC099326	AC099326	Homo sapi
	41	27	2.0	153448	9	AC063923	AC063923	Homo sapi
	42	27	2.0	174531	9	AC104449	AC104449	Homo sapi
	43	27	2.0	177672	2	AC010959	AC010959	Homo sapi
	44	27	2.0	179585	9	AC106783	AC106783	Homo sapi
	45	27	2.0	189013	6	AX706971	AX706971	Sequence
	46	27	2.0	189013	6	AX707901	AX707901	Sequence
	47	27	2.0	189032	9	AC025277	AC025277	Homo sapi
	48	27	2.0	199045	2	AC110206	AC110206	Mus muscu
	49	27	2.0	204855	10	AC133100	AC133100	Mus muscu
	50	27	2.0	232802	2	AC099108	AC099108	Rattus no
	51	26	2.0	651	5	AB033881	AB033881	Coturnix
	52	26	2.0	698	8	BO1FAB2	BO1FAB2	B.oleracea
	53	26	2.0	913	10	BC061227	BC061227	Mus muscu
	54	26	2.0	931	10	BC049725	BC049725	Mus muscu
	55	26	2.0	1077	9	BC062787	BC062787	Homo sapi
	56	26	2.0	1078	8	AY173062	AY173062	Chrysanth
	57	26	2.0	1186	10	BC063154	BC063154	Rattus no
	58	26	2.0	1238	3	AK115014	AK115014	Ciona int
	59	26	2.0	1570	10	BC049624	BC049624	Mus muscu
	60	26	2.0	1888	5	BC056828	BC056828	Danio rer
	61	26	2.0	2250	10	BC059877	BC059877	Mus muscu
	62	26	2.0	2619	10	BC019375	BC019375	Mus muscu
	63	26	2.0	2893	10	BC042763	BC042763	Mus muscu
	64	26	2.0	2955	10	BC021442	BC021442	Mus muscu
	65	26	2.0	3270	9	BC050552	BC050552	Homo sapi

66	26	2.0	4059	10	BC056981	Mus muscu
67	26	2.0	4798	10	BC056365	Mus muscu
68	26	2.0	49830	6	AX059541	Sequence
C 69	26	2.0	78239	8	AB012241	Arabidops
C 70	26	2.0	86424	2	AC010222	Homo sapi
C 71	26	2.0	96685	8	AC007087	Arabidops
72	26	2.0	115435	9	AC097626	Homo sapi
C 73	26	2.0	130965	2	AC026773	Homo sapi
C 74	26	2.0	142290	9	AC102011	Mus muscu
C 75	26	2.0	142290	9	AC102011	Mus muscu
C 76	26	2.0	145453	8	AC012477	Genomic S
C 77	26	2.0	148043	9	AC008541	Homo sapi
C 78	26	2.0	157848	9	AC0079763	Homo sapi
C 79	26	2.0	174645	2	AC136857	Rattus no
C 80	26	2.0	183147	8	AC012392	Genomic S
C 81	26	2.0	184132	2	AC120345	Mus muscu
82	26	2.0	187162	2	AC139042	Mus muscu
83	26	2.0	192309	2	BX545915	Danio rer
84	26	2.0	192586	10	AL645972	Mouse DNA
C 85	26	2.0	195208	10	AL591146	Mouse DNA
C 86	26	2.0	199662	2	AC131241	Mus muscu
C 87	26	2.0	199987	8	ATCHRIV15	Arabidops
C 88	26	2.0	201174	10	AC113953	Mus muscu
C 89	26	2.0	207136	10	AC122268	Mus muscu
90	26	2.0	211393	2	AC119244	Mus muscu
91	26	2.0	211803	2	AC117813	Mus muscu
92	26	2.0	219288	2	AC094180	Rattus no
C 93	26	2.0	243868	2	AC094180	Rattus no
C 94	26	2.0	245267	2	AC106568	Rattus no
95	26	2.0	255156	2	AC111130	Mus muscu
96	26	2.0	260876	2	AC127709	Rattus no
97	26	2.0	270618	2	AC133365	Rattus no
C 98	26	2.0	312459	2	AL928900	Danio rer
C 99	25	1.9	62	9	S52152831	Sequence
C 100	25	1.9	191	6	AX332084	Sequence
C 101	25	1.9	191	6	AX333538	Sequence
C 102	25	1.9	191	6	AX407878	Sequence
C 103	25	1.9	222	6	AX407940	Sequence
C 104	25	1.9	252	3	AF215118	Sequence
C 105	25	1.9	330	6	AF251319	Macaca mu
C 106	25	1.9	365	9	AF251319	Macaca mu
C 107	25	1.9	390	6	AF645775	Sequence
C 108	25	1.9	430	6	BC058034	Homo sapi
C 109	25	1.9	438	11	H8542C5S	Sequence
C 110	25	1.9	444	10	BC059153	Rattus no
C 111	25	1.9	452	9	AF487338	Homo sapi
C 112	25	1.9	469	6	AR019336	Sequence
C 113	25	1.9	469	6	AR028949	Sequence
C 114	25	1.9	469	6	AR037083	Sequence
C 115	25	1.9	469	6	AR049639	Sequence
C 116	25	1.9	469	6	AR145297	Sequence
C 117	25	1.9	538	9	BC003669	Homo sapi
C 118	25	1.9	540	5	AY168623	Danio rer
C 119	25	1.9	612	9	BC056906	Homo sapi
C 120	25	1.9	657	6	AX525532	Sequence
C 121	25	1.9	671	10	BC027578	Mus muscu
C 122	25	1.9	704	11	BV023499	S212P6409
C 123	25	1.9	716	6	BD216887	Novel hum
C 124	25	1.9	729	8	ATAPT	A.thaliana
C 125	25	1.9	748	6	BD019727	Novel gen
C 126	25	1.9	748	6	BD099665	Novel gen
C 127	25	1.9	775	11	BV070662	S212P6049
C 128	25	1.9	780	11	BV038342	S212P6390
C 129	25	1.9	836	3	AF457549	Anopheles
C 130	25	1.9	876	10	BC008275	Mus muscu
C 131	25	1.9	893	9	BC009228	Homo sapi
C 132	25	1.9	938	3	AB072733	Hemicentr
C 133	25	1.9	953	9	BC011751	Homo sapi
C 134	25	1.9	1073	10	BC049992	Mus muscu
C 135	25	1.9	1091	9	BC000566	Homo sapi
C 136	25	1.9	1149	10	BC022645	Mus muscu
C 137	25	1.9	1171	9	HSM800589	Homo sapi
C 138	25	1.9	1182	10	BC049567	Mus muscu

ALIGNMENTS

139	25	1.9	1214	9	AK026312	
140	25	1.9	1226	9	BC039664	Homo sapi
141	25	1.9	1230	10	BC060975	Mus muscu
142	25	1.9	1238	10	BC052530	Mus muscu
143	25	1.9	1278	5	AF100931	Oncohync
144	25	1.9	1280	17	AF130066	Homo sapi
145	25	1.9	1299	9	BC026126	Homo sapi
146	25	1.9	1330	9	BC012495	Homo sapi
147	25	1.9	1357	9	BC020805	Homo sapi
148	25	1.9	1417	9	BC029650	Homo sapi
149	25	1.9	1483	5	AY120892	Danio rer
150	25	1.9	1502	9	BC008376	Homo sapi
RESULT 1						
LOCUS	BC001222		3259 bp	mRNA	linear	PRI 04-OCT-2003
DEFINITION	Homo sapiens hypothetical protein FLJ10759, mRNA (cDNA clone MGC:938 IMAGE:3355572), complete cds.					
ACCESSION	BC001222					
VERSION	BC001222.1		GI:12654758			
KEYWORDS	MGC.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
AUTHORS	1 (bases 1 to 3259)					
	Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raja, S.S., Joquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, M.A., Schnerch, A., Schein, J.E., Jones, S.J. and Matra, M.A.					
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences					
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	99	(26)	16899-16903	(2002)	
MEDLINE	22388257					
PUBMED	12477932					
REFERENCE	2 (bases 1 to 3259)					
AUTHORS	Strausberg, R.					
TITLE	Direct Submission					
JOURNAL	Submitted (11-DEC-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA					
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov					
COMMENT	Contact: MGC help desk Email: cgabs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) BC Cancer Agency, Vancouver, BC, Canada info@bcsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Sco Len Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven					

Ness, Pawan Pandoh, Anna-Jiisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 6 Row: J Column: 17
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 8922647.

FEATURES

Location/Qualifiers
1..3259
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:938 IMAGE:3355572"
/tissue_types="Eye, Retinoblastoma"
/clone_lib="NIH MGC_16"
/lab_host="DH10B-R"
/note="vector: pOTB7"
1..3259
/gene="FLJ10759"
/db_xref="LocusID:55223"
34..1461
/codon_start=1
/product="hypothetical protein FLJ10759"
/protein_id="AAH01222.1"
/db_xref="GI:12654759"
/db_xref="LocusID:55223"
/translation="MACS1KDELLCSISYQDPVSLGCEHYFCRRCTEHWVROEA QGARDPECBRTAFAPALPSLKLANIVERYSFFPLDALNARRAAPQOAHDKVLF CLTDRAILLCFCDEPALHQHQTGDDAFDELQELKQALQDQSERHEALOLL KRLAETKSTQSLRTTIGAEFRHLRLREROKAMLELEADTLTDIEQVORY SOQLRVQGAQILQERLADTRHTFLAGVASLSERLKGIHETNLTYEDPFTSKYTG PLYQTWKSIFQDHPVPALATDPCTAQRLILSDCTIYVGNHLPQLQDQPKRF DVEVSVLGSFSSGHHYEVVVAETQWVIGLAHEAARSKGSIQIOPSRGFYIVMH DGNQYSACTEPTRLNRDLKDVGVFLDYDQGLLIFYNADMSWLYTFRKPPGKLC SYSPGCGSHANGKNVQPLRINTVRI"
64..192
/note="RING; Region: Ring finger"
/db_xref="CDD:smart00184"
445..798
/note="BBC; Region: B-Box C-terminal domain"
/db_xref="CDD:smart00502"
913..1074
/note="PRY; Region: associated with SPRY domains"
1075..1413
/note="SPRY; Region: Domain in SPla and the Ryanodine Receptor."
/db_xref="CDD:smart00449"
ORIGIN
Query Match 100.0%; Score 1327; DB 9; Length 3259;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCACAGTGGTACAGGTAGTACCTGGTCTAGGTTGCTGAGAGCCCAACCTCTCCTGC 60
Db 1928 TCACAGTGGTACAGGTAGTACCTGGTCTAGGTTGCTGAGAGCCCAACCTCTCCTGC 1987
QY 61 CACCCCCACACCAAGAACATATAGTTCTCTCCACTGATCTGCTGGTCAAGTAT 120
Db 1988 CACCCCCACACCAAGAACATATAGTTCTCTCCACTGATCTGCTGGTCAAGTAT 2047
QY 121 GATGCTGGCTGTGGAGGACCTGGTGTAGTTCACACATATAGTCAATGTGCCA 180
Db 2048 GATGCTGGCTGTGGAGGACCTGGTGTAGTTCACACATATAGTCAATGTGCCA 2107
QY 181 CCACCTTCTGCCCCACAGCCGAGGACAGGTGAGGTATACCCAAAGCTGATCCAG 240
Db 2108 CCACCTTCTGCCCCACAGCCGAGGACAGGTGAGGTATACCCAAAGCTGATCCAG 2167

Db	3248	AAAAAA	3254		
RESULT 2					
BC007999			3259 bp	mRNA	linear
LOCUS					PRI 03-OCT-2003
DEFINITION					Homo sapiens hypothetical protein FLJ10759, mRNA (cDNA clone
					MGC:15757 IMAGE:3357436), complete cds.
ACCESSION					BC007999
VERSION					BC007999.1
KEYWORDS					GI:14124949
SOURCE					MGC.
ORGANISM					Homo sapiens (human)
REFERENCE					
AUTHORS					
					Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
					Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
					1 (bases 1 to 3259)
					Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
					Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
					Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
					Hopkins, R.F., Jordan, H., Moore, I., Max, S.I., Wang, J., Hsieh, F.,
					Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
					Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
					Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
					Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
					Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
					McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
					Worley, K.C., Hale, S., Garcia, A.M., Gay, D.J., Hulyk, S.W.,
					Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
					Fabey, J., Helton, E., Kettman, M., Madan, A.C., Rodrigues, S.,
					Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
					Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
					Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
					Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalilus, D.E.,
					Schneerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
					Generation and initial analysis of more than 15,000 full-length
					human and mouse cDNA sequences
					Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
					22388257
					12477932
					2 (bases 1 to 3259)
					Strausberg, R.
					Direct Submission
					Submitted (15-MAY-2001) National Institutes of Health, Mammalian
					Gene Collection (MGC), Cancer Genomics Office, National Cancer
					Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
					USA
					NIH-MGC Project URL: http://mgc.nci.nih.gov
					Contact: MGC help desk
					Email: cgabbs-r@mail.nih.gov
					Tissue Procurement: AICC
					cDNA Library Preparation: Rubin Laboratory
					cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
					DNA Sequencing by: Genome Sequence Centre,
					BC Cancer Agency, Vancouver, BC, Canada
					info@bcgsc.bc.ca
					Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
					Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
					Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
					Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
					Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline
					Schein, Duane Smalilus, Michael Smith, Lorraine Spence, Jeff Stott,
					Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
					George Yang, Scott Zuyderduyn, Marco Marra.
					Clone distribution: MGC clone distribution information can be found
					through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
					Series: IRAL Plate: 23 Row: a Column: 20.
					Location/Qualifiers
					1. .3259
					/organism="Homo sapiens"
					/mol_type="mRNA"
					/db_xref="taxon:9606"
FEATURES					
source					

gene	1. .3259	/clone="MGC:15757 IMAGE:3357436"
		/tissue_type="Eye, retinoblastoma"
		/clone_lib="NIH MGC 16"
		/lab_host="DH10B-R"
		/notes="Vector: pOTB7"
CDS	34. .1461	/genes="FLJ10759"
		/db_xref="LocusID:55223"
		/codon_start=1
		/product="hypothetical protein FLJ10759"
		/protein_id="AAH07999.1"
		/db_xref="GI:14124950"
		/db_xref="LocusID:55223"
		/translation="MACSKDELCSICLSIYQDPVSLGCEHYFCRCRITHEHWROEA
		QAGRDCEGRTEFAEPALPSKLNIIVERYSPFLDAILNARRAAPCOAHDKVLF
		CLTDALLCFCEDEPALHBOHTGDDAFDELQRELQLOLQADSERHETALQLL
		KROLAETKSKSLRTIIGAPERHLERLPEROKMELEFARTPLTDEOKVQRY
		SOQRKVQGAQILQERLAETDRHTFLAGVASERLKGKHETNLTYEDPPTSKYR
		PLQYTIWKSLEFQDHPVPAALTDPGTAHQRLISDDCTIIVAGNLHPQLQDSKRF
		DVGVSIGSEAFSSGVHYWEVVAEKQWVIGLAHEAASRSGSIQIPSGFYCIVNH
		DGNOYSACTEPTRLNVRDKLVGVFLDYDQGLLIPYNADDSWLYTFREKPPGKLC
		SYFSPQSHANGNVQLRINTVRI"
misc_feature	64. .192	/notes="RING; Region: Ring finger"
		/db_xref="CDD:smart00184"
misc_feature	445. .798	/note="BBC; Region: B-Box C-terminal domain"
		/db_xref="CDD:smart00502"
misc_feature	913. .1074	/note="PRV; Region: associated with SPRY domains"
		/db_xref="CDD:smart00589"
misc_feature	1075. .1413	/note="SPRY; Region: Domain in SPLa and the Ryanodine
		Receptor."
		/db_xref="CDD:smart00449"
ORIGIN		
	Query Match	100.0%; Score 1327; DB 9; Length 3259;
	Best Local Similarity	100.0%; Pred. No. 0;
	Matches 1327; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1	TCCACAGTGTGCACAGGTAGTACCTGTCCTAGGTTGCCTGAGAGCAACCTCCCTGC 60
Db	1928	TCCACAGTGTGCACAGGTAGTACCTGTCCTAGGTTGCCTGAGAGCAACCTCCCTGC 1987
QY	61	CACCCCCACACCAAGAACTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Db	1988	CACCCCCACACCAAGAACTATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2047
QY	121	GATGCTGTGGCTGTGGAAGGCACCTGGTGTGAGTCCACACATATAGTCATGTGCA 180
Db	2048	GATGCTGTGGCTGTGGAAGGCACCTGGTGTGAGTCCACACATATAGTCATGTGCA 2107
QY	181	CCACCTTCTCTCCACAGCCGAGGACAGGTGAGGTGAGGTATACCCAAAGCTGATCAGAG 240
Db	2108	CCACCTTCTCTCCACAGCCGAGGACAGGTGAGGTGAGGTATACCCAAAGCTGATCAGAG 2167
QY	241	CCATTAGCTTAAAGCAACTGCAGGACAAAGCTCTCTGGATGATCGAGGTCCCGAGTAG 300
Db	2168	CCATTAGCTTAAAGCAACTGCAGGACAAAGCTCTCTGGATGATCGAGGTCCCGAGTAG 2227
QY	301	CTCTGAACAAGAGTCCAGCAACCCCTCTTCCAGCCAGGCTCTGTGACCTGTAGGTGCA 360
Db	2228	CTCTGAACAAGAGTCCAGCAACCCCTCTTCCAGCCAGGCTCTGTGACCTGTAGGTGCA 2287
QY	361	GGAGGCTTCCAGAGCAGTGTGTGTAATTAGGACCAACCACTGGGAGGGCTGTGGCT 420
Db	2288	GGAGGCTTCCAGAGCAGTGTGTGTAATTAGGACCAACCACTGGGAGGGCTGTGGCT 2347
QY	421	AGACCCCTTGTGACACTTGGCATCTATCTCAGTGTAGGATCCTGCTGCAGAAAAAAGAGC 480
Db	2348	AGACCCCTTGTGACACTTGGCATCTATCTCAGTGTAGGATCCTGCTGCAGAAAAAAGAGC 2407

QY 481 CACCTGTAGCTGGTTTAATTAGACAAGGATTTTACTACCTGGCCCTGGTGGCTTGCAGAAA 540
Db |||||
QY 2408 CACCTGTAGCTGGTTTAATTAGACAAGGATTTTACTACCTGGCCCTGGTGGCTTGCAGAAA 2467
Db |||||
QY 541 TTGTTGGAAGAGCTGGAGAGCAGACTCTGCTGAATTTCCAGGAATCCCGAGCCAGAT 600
Db |||||
QY 2468 TTGTTGGAAGAGCTGGAGAGCAGACTCTGCTGAATTTCCAGGAATCCCGAGCCAGAT 2527
Db |||||
QY 601 TCATCATGCTCTGTTGTGACCAAGAAAGCTGCCCCATCTGCAGGAGCACTATGCCGAGA 660
Db |||||
QY 2528 TCATCATGCTCTGTTGTGACCAAGAAAGCTGCCCCATCTGCAGGAGCACTATGCCGAGA 2587
Db |||||
QY 661 AAGCTGCTGACTGCAGAACTAGGTCCTCTGCCACGGTCCGTGCCAGCCCAATAGATGTC 720
Db |||||
QY 2588 AAGCTGCTGACTGCAGAACTAGGTCCTCTGCCACGGTCCGTGCCAGCCCAATAGATGTC 2647
Db |||||
QY 721 CTGAGGCTGCTGCCCTCTCCCACTTCACTCAGTCTCCCAATCTAAATTTTCAAGAGATT 780
Db |||||
QY 2648 CTGAGGCTGCTGCCCTCTCCCACTTCACTCAGTCTCCCAATCTAAATTTTCAAGAGATT 2707
Db |||||
QY 781 CTGTTTGGGGAATTAAGTCAGATCCAGAACCTTGGCTGCAAGGGATCTGGGAATGT 840
Db |||||
QY 2708 CTGTTTGGGGAATTAAGTCAGATCCAGAACCTTGGCTGCAAGGGATCTGGGAATGT 2767
Db |||||
QY 841 CATTTCCCTAGAGGAAGTTAGGTTGGGTGGAGCAAGCCCACTGGCTTTTCTGCCAC 900
Db |||||
QY 2768 CATTTCCCTAGAGGAAGTTAGGTTGGGTGGAGCAAGCCCACTGGCTTTTCTGCCAC 2827
Db |||||
QY 901 AGCATCCAATCGTGAAGAACTCGGAGAGGGTGAGTCCACATCTAGGGTTGCTCTGCC 960
Db |||||
QY 2828 AGCATCCAATCGTGAAGAACTCGGAGAGGGTGAGTCCACATCTAGGGTTGCTCTGCC 2887
Db |||||
QY 961 CTTGGCTCTATCCCTGCCAGAGTGGGAACTGGAGAGTGGGCTGCAAGACTGAGCCTA 1020
Db |||||
QY 2888 CTTGGCTCTATCCCTGCCAGAGTGGGAACTGGAGAGTGGGCTGCAAGACTGAGCCTA 2947
Db |||||
QY 1021 AATGTCCTCCCGGCTTGACTTTCTTTCTAGTCTGGGCGCTAGATTCTGCATTTGGGG 1080
Db |||||
QY 2948 AATGTCCTCCCGGCTTGACTTTCTTTCTAGTCTGGGCGCTAGATTCTGCATTTGGGG 3007
Db |||||
QY 1081 TCTCTGACACACACACACATCCCAAGTAGCCGAGAGAGTAAACACAGGGGTTCTTAA 1140
Db |||||
QY 3008 TCTCTGACACACACACACATCCCAAGTAGCCGAGAGAGTAAACACAGGGGTTCTTAA 3067
Db |||||
QY 1141 AATGGCTGCCCGCCACCCGCGCTCCCTTGGGCAAGAGGATTTGTCAGCCCTACCCCA 1200
Db |||||
QY 3068 AATGGCTGCCCGCCACCCGCGCTCCCTTGGGCAAGAGGATTTGTCAGCCCTACCCCA 3127
Db |||||
QY 1201 AGCCTTCACTACAGAACTCTGGGCGCCACCCGAGTATTTTATTTAAATGTTGCCCA 1260
Db |||||
QY 3128 ACCCTTCACTACAGAACTCTGGGCGCCACCCGAGTATTTTATTTAAATGTTGCCCA 3187
Db |||||
QY 1261 TTTTATGAGTTATGATCAATTTGATTAATTAAGTTACAGATGTCAGGCAAAAAA 1320
Db |||||
QY 3188 TTTTATGAGTTATGATCAATTTGATTAATTAAGTTACAGATGTCAGGCAAAAAA 3247
Db |||||
QY 1321 AAAAAAA 1327
Db |||||
QY 3248 AAAAAAA 3254
Db |||||

RESULT 3
BC011689
LOCUS
DEFINITION Homo sapiens hypothetical protein FLJ10759, mRNA (cdna clone 3259 bp mRNA linear PRI 19-DEC-2003
MGC:19672 IMAGE:3353034), complete cds.
ACCESSION BC011689
VERSION BC011689.2 GI:40225630
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1. (bases 1 to 3259)
Strausberg, R., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, D.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2. (bases 1 to 3259)
Strausberg, R.
Direct Submission
Submitted (30-JUL-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Dec 19, 2003 this sequence version replaced gi:15079757.
Contact: MGC help desk
Email: cgabs@email.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lario, P., Legaspi, R.,
Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 27 Row: b Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.

FEATURES
source

1. 3259
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:19672 IMAGE:3353034"
/tissue type="Eye, retinoblastoma"
/clone_lib="NIH MGC_16"
/lab_host="DH10B-R"
/note="vector: pOTB7"
1. 3259
/gene="FLJ10759"
/db_xref="LocusID:55223"
34_1461
/gene="FLJ10759"
/codon_start=1

```

/product="hypothetical protein FLJ10759"
/protein_id="AAH11689.1"
/db_xref="GI:15079758"
/db_xref="LocusID:55223"
/translation="MACSKDELCSILSIYODPVSLGCEHYFCRRCTEHWVROEA
QGARDPECRRTAPALAPSLXLANIVERYSPFLDAILNARRAAPPCOAHDKVKLF
CLDRALLCFDEPAPALHOQVTGDDAFDELQRELKOLOLQALOSEREHTEALOLL
KRGIAETKSLRTTIGFAELHLRLREROKAMLELEANTARLIDIEGKQRY
SQLRKVQEQALQERLAEDRHTFLAGVASLSERUKGKHETNLTDFPFKSTIG
PLOTWVLSFQDHPVPAALTLDPGTAHORLISDDCTIVAYGNLHPQLQDSPKRF
DVSVYVGSFSGVHWBVAEKOTWVIGLAHEAASRKSIOIQPSRGFYCIYVH
DGNQYSACTPEPFLNVRDKLVGVFLDYDQGLLIPYNADMSWLYTFREKPPGKLC
SYFSPGQSHANGKVPRLNVR1"
misc_feature
64..204
/gene="FLJ10759"
/note="RING; Region: RING-finger (Really Interesting New
Gene) domain, a specialized type of Zn-finger of 40 to 60
residues that binds two atoms of zinc"
/db_xref="CDD:cd00162"
misc_feature
445..798
/gene="FLJ10759"
/note="B-Box C-terminal domain"
/db_xref="CDD:smart00502"
misc_feature
913..1074
/gene="FLJ10759"
/note="PRY; Region: associated with SPRY domains"
/db_xref="CDD:smart00589"
misc_feature
1075..1413
/gene="FLJ10759"
/note="SPRY; Region: Domain in Spla and the Ryanodine
Receptor."
/db_xref="CDD:smart00449"

ORIGIN
Query Match 100.0%; Score 1327; DB 9; Length 3259;
Best local Similarity 100.0%; Pred. No. 0;
Matches 1327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCACAGTGGTCCACAGGTAGTACCTGGTCTAGGTTGGCTGAGAGCCACCTCTCCG 60
DB 1928 TCCACAGTGGTCCACAGGTAGTACCTGGTCTAGGTTGGCTGAGAGCCACCTCTCC 1987

QY 61 CACCCCCACACAGAGTATATGTTCTCTCCACAGTATCTGGTCTAGTAT 120
DB 1988 CACCCCCACACAGAGTATATGTTCTCTCCACAGTATCTGGTCTAGTAT 2047

QY 121 GATGCTGTGGCTGTGGAGGACCTGTGTAGTTCACACATATATAGTCAATG 180
DB 2048 GATGCTGTGGCTGTGGAGGACCTGTGTAGTTCACACATATATAGTCAATG 2107

QY 181 CCACCTTCCTGCCACAGCCGAGGACAGGGTGGGTATATCCAAAGCTGATG 240
DB 2108 CCACCTTCCTGCCACAGCCGAGGACAGGGTGGGTATATCCAAAGCTGATG 2167

QY 241 CCATATAGCTTAAAGCACTGAGCAGACAGCTCCCTGATGATGAGTCCCA 300
DB 2168 CCATATAGCTTAAAGCACTGAGCAGACAGCTCCCTGATGATGAGTCCCA 2227

QY 301 CTCTGAACAGAGTCCAGCCAAACCTCTTCCAGCCAGGCTCTGTGACCTG 360
DB 2228 CTCTGAACAGAGTCCAGCCAAACCTCTTCCAGCCAGGCTCTGTGACCTG 2287

QY 361 GGAGGCTTCAGAGCAGTGTGTGTAATAGGACCCAGCAGTGGAGGGGCTGT 420
DB 2288 GGAGGCTTCAGAGCAGTGTGTGTAATAGGACCCAGCAGTGGAGGGGCTGT 2347

QY 421 AGACCCCTTGTACAGCTTGGCATCTATCTCAGTATAGTCTCTGTGAC 480
DB 2348 AGACCCCTTGTACAGCTTGGCATCTATCTCAGTATAGTCTCTGTGAC 2407

QY 481 CACTGTAGCTGGTGTATATAGACAGAGTATCTACCTGGCCCTCGTGGCT 540
DB 2408 CACTGTAGCTGGTGTATATAGACAGAGTATCTACCTGGCCCTCGTGGCT 2467

```

```

QY 541 TTGTTGGAAGAGCTGGAGAGCAGACTCTCTGTAATTTCCAGGAATCCCGAGCCAGAT 600
DB 2468 TTGTTGGAAGAGCTGGAGAGCAGACTCTCTGTAATTTCCAGGAATCCCGAGCCAGAT 2527

QY 601 TCATCATGCTGTGTGTGACAGAAAGCTGCCCCCATCTGTCAGAGAGCCACTATGCCAGA 660
DB 2528 TCATCATGCTGTGTGTGACAGAAAGCTGCCCCCATCTGTCAGAGAGCCACTATGCCAGA 2587

QY 661 AAGCTGCTGACTCGAGAACTTAGCTCCCTCTGTCACAGGTCGTCGCCAGCAATAGATGTC 720
DB 2588 AAGCTGCTGACTCGAGAACTTAGCTCCCTCTGTCACAGGTCGTCGCCAGCAATAGATGTC 2647

QY 721 CTGAGGCTGCCCTCTCCCACTTCACTAGTTTCCCAAAATCTAAATTTTACAGAGATT 780
DB 2648 CTGAGGCTGCCCTCTCCCACTTCACTAGTTTCCCAAAATCTAAATTTTACAGAGATT 2707

QY 781 CTGTTGGGGAACTTAAGTCAGATCCAGAACTTGGCTGCAAGGGAGTCTGGGAAATGT 840
DB 2708 CTGTTGGGGAACTTAAGTCAGATCCAGAACTTGGCTGCAAGGGAGTCTGGGAAATGT 2767

QY 841 CATTTCCCTAGAGAAAGTTAGGTTGGTGGAGCAAGCCACCTGGCTTTTCTGCGCAC 900
DB 2768 CATTTCCCTAGAGAAAGTTAGGTTGGTGGAGCAAGCCACCTGGCTTTTCTGCGCAC 2827

QY 901 AGCATCCAATCGTGAAGAACTCGGGAGAGGTGAGTCCACATCTAGGTTGTCTGCCC 960
DB 2828 AGCATCCAATCGTGAAGAACTCGGGAGAGGTGAGTCCACATCTAGGTTGTCTGCCC 2887

QY 961 CTTGGCTTATCCCTGCCAGAGTGGGAAGTGGAGTGGCTGCAAGACTGAGCCCTA 1020
DB 2888 CTTGGCTTATCCCTGCCAGAGTGGGAAGTGGAGTGGCTGCAAGACTGAGCCCTA 2947

QY 1021 AATGTCCTCCGCGCTTGACTTTTCTTCTAGTCTCTGGGCTAGATTCTGCACTTGGG 1080
DB 2948 AATGTCCTCCGCGCTTGACTTTTCTTCTAGTCTCTGGGCTAGATTCTGCACTTGGG 3007

QY 1081 TCTCTGACACACACACCATCCCAAAGTAGCGGAGAGCTTAAACACAGGGGGTCTTAA 1140
DB 3008 TCTCTGACACACACACCATCCCAAAGTAGCGGAGAGCTTAAACACAGGGGGTCTTAA 3067

QY 1141 AATGCTGCTCCCGCCACCCGGGCTCCCTTTGGGCAAAAGGAATTTGTGAGCCCTACCCCA 1200
DB 3068 AATGCTGCTCCCGCCACCCGGGCTCCCTTTGGGCAAAAGGAATTTGTGAGCCCTACCCCA 3127

QY 1201 ACCCTTCAACTACCAAGATCTGGGCCACCCAGCAGTATTTTATTTAAATGTTGCCCA 1260
DB 3128 ACCCTTCAACTACCAAGATCTGGGCCACCCAGCAGTATTTTATTTAAATGTTGCCCA 3187

QY 1261 TTTTATGAGTTATGATCAATTTGATTAATTAAGTTACAGATGTCAAAAA 1320
DB 3188 TTTTATGAGTTATGATCAATTTGATTAATTAAGTTACAGATGTCAAAAA 3247

QY 1321 AAAAAA 1327
DB 3248 AAAAAA 3254

```

```

RESULT 4
BC012152 3259 bp mRNA linear PRI 04-OCT-2003
Homo sapiens hypothetical protein FLJ10759, mRNA (cdna clone
LOCUS MGC:20370 IMAGE:4558639), complete cds.
DEFINITION
ACCESSION BC012152
VERSION BC012152.1 GI:15082475
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3259)
AUTHORS Strausberg R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,

```

Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P., Datchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Scapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McSwain,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahy,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.A., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shcherchenko,Y., Bouffard,G.C., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywiński,M.I., Skalski,U., Smalius,D.E., Schnerch,A., Schein,J.E., Jones,S.J., and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 3259)

Strausberg,R.

Direct Submission

Submitted (02-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs.romail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywiński, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline Schein, Duane Smalius, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuydam, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 29 Row: d Column: 2
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 8922647.
tactin-10

FEATURES	SOURCE
----------	--------

source

```

1. .3259
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:20370 IMAGE:4558639"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH MGC_16"
/lab_host="DH10B-R"
/note="Vector: pOTR7"
1. .3259
/gene="FLJ10759"
/db_xref="LocusID:55223"
34. .1461
/codon_start=1

```

gene

```
/note="vector: pOTB7"
1. 3259
/gene="FLJ10759"
/db_xref="LocusID:55223"
34.1461
/codon start=1
```

CDS

```

/ww_xref="LocusID:55223"
34..1461
/codon start=1

```

misc_feature	64..192	/note="KING; Region: Ring finger" /db_xref="CDD:smart00184"
misc_feature	445..798	/note="BBC; Region: B-Box C-terminal domain" /db_xref="CDD:smart00502"
misc_feature	913..1074	/note="PRY; Region: associated with SPRY domains" /db_xref="CDD:smart00589"
misc_feature	1075..1413	/note="SPRY; Region: Domain in Spla and the Ryanodine Receptor." /db_xref="CDD:smart00449"
ORIGIN		
Query Match	100.0%;	Score 1327; DB 9; Length 3259;
Best Local Similarity	100.0%;	Pred. No. 0;
Matches 1327;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	TCCACAGTGGTCACAGTAGTACTGTGGTCCTAGGGTTGCTCGAGAGCCAACTCTCCTGC 60
Db	1928	TCCACAGTGGTCACAGTAGTACTGTGGTCCTAGGGTTGCTCGAGAGCCAACTCTCCTGC 1987
Qy	61	CACCCCCACACCAAGAACTATATGGTTCCTACTCTCCACACTGATCTGCTGTCAGTGAT 120
Db	1988	CACCCCCACACCAAGAACTATATGGTTCCTACTCTCCACACTGATCTGCTGTCAGTGAT 2047
Qy	121	GATGCTGTGGCTGTGGAAGGCCACCTGGTAGTTGAGTCACACATATTAGTCATCTGCCA 180
Db	2048	GATGCTGTGGCTGTGGAAGGCCACCTGGTAGTTGAGTCACACATATTAGTCATCTGCCA 2107
Qy	181	CCACCTTCCTGCCACAGGCCGAGGACAGGGTGAGGHPATACCAAAGCTGATGCAGAG 240
Db	2108	CCACCTTCCTGCCACAGGCCGAGGACAGGGTGAGGHPATACCAAAGCTGATGCAGAG 2167
Qy	241	CCCATTAGCCTTAAAGCAACTGCAGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 300
Db	2168	CCCATTAGCCTTAAAGCAACTGCAGACAAGCCTCCCTGGATGATCGAGGTCCCCAGTAG 2227
Qy	301	CTCTGACACAGAGTCAGACCAACCTCTTCAGCCAGGCTCTGTGACCTGCTAGGGTGCA 360
Db	2228	CTCTGACACAGAGTCAGACCAACCTCTTCAGCCAGGCTCTGTGACCTGCTAGGGTGCA 2287
Qy	361	GGAGGGCTTCAGAAAGCAGTTGTTGAATTAGACCCACAGCACTGGAGGGCTGTGGCT 420
Db	2288	GGAGGGCTTCAGAAAGCAGTTGTTGAATTAGACCCACAGCACTGGAGGGCTGTGGCT 2347
Qy	421	AGACCCCTTGTGAGACTTGGCATCTATCTCAGTTAGGATCCCTGCTGCAGAAAAAAGAGC 480
Db	2348	AGACCCCTTGTGAGACTTGGCATCTATCTCAGTTAGGATCCCTGCTGCAGAAAAAAGAGC 2407
Qy	481	CACCTGTAGCTGGTTTAATTAGACAGGATTTACTACTGGCCCCCTGGTGGCTTGCAAAA 540
Db	2408	CACCTGTAGCTGGTTTAATTAGACAGGATTTACTACTGGCCCCCTGGTGGCTTGCAAAA 2467
Qy	541	TTGTTGGAGAGCTGGAGAGCAGACTCTGCTGAATTTCCAGGAATCTCCAGGCCAGAT 600
Db	2468	TTGTTGGAGAGCTGGAGAGCAGACTCTGCTGAATTTCCAGGAATCTCCAGGCCAGAT 2527
Qy	601	TCATCATGTCGTGTGTGACCAAGAAAGCTGCCCCATCTGCGAGAAAGCACTATGCGAGA 660
Db	2528	TCATCATGTCGTGTGTGACCAAGAAAGCTGCCCCATCTGCGAGAAAGCACTATGCGAGA 2587
Qy	661	AAAGCTGCTGACTCGAGAACTAGGCTCCCTCTGCGACAGCGTCCGTCAGAGCAATAGATGTC 720
Db	2588	AAAGCTGCTGACTCGAGAACTAGGCTCCCTCTGCGACAGCGTCCGTCAGAGCAATAGATGTC 2647

QY	721	CTGAGGCCTGCCCTCTCCCACTTACACTCAGTTCCCAATCTAATAATTTTTACAAGAGATT	780
Db	2648	CTGAGGCCTGCCCTCTCCCACTTACACTCAGTTCCCAATCTAATAATTTTTACAAGAGATT	2707
QY	781	CTGTGTTGGGGAACTTAAGTCAGATCCAGAACCTTGCGTGCAGGAGAGTCTGGGAATGT	840
Db	2708	CTGTGTTGGGGAACTTAAGTCAGATCCAGAACCTTGCGTGCAGGAGAGTCTGGGAATGT	2767
QY	841	CATTCCCTAGAGAAGATTAGGTTGGGTGGAGAACCCCACCCTGCGTTTTCTGCCAC	900
Db	2768	CATTCCCTAGAGAAGATTAGGTTGGGTGGAGAACCCCACCCTGCGTTTTCTGCCAC	2827
QY	901	AGCATCCAATCGTAGAACTCGGAGAGGTTGGAGTCCACATCTAGGGTTGTCCTGCC	960
Db	2828	AGCATCCAATCGTAGAACTCGGAGAGGTTGGAGTCCACATCTAGGGTTGTCCTGCC	2887
QY	961	CTTGCGTCTATCCCTGCCAGAGTGGAACTGGAGAGTGGGCTCCAGAATGAGCCTTA	1020
Db	2888	CTTGCGTCTATCCCTGCCAGAGTGGAACTGGAGAGTGGGCTCCAGAATGAGCCTTA	2947
QY	1021	AATGTCCTCCCGGCTTGACTTTCTTTCTAGTCCCTGGGCTTAGATCTGACATCTAGGG	1080
Db	2948	AATGTCCTCCCGGCTTGACTTTCTTTCTAGTCCCTGGGCTTAGATCTGACATCTAGGG	3007
QY	1081	TCTCTGACAAACACACATCCCAAAGTAGCCGAGAGCTAAACACAGGGGTTCTTAA	1140
Db	3008	TCTCTGACAAACACACATCCCAAAGTAGCCGAGAGCTAAACACAGGGGTTCTTAA	3067
QY	1141	AATGGCTGCCCGCCACCCCGGGCTCCCTTGGGCAAAAGAAATGTCAGCCCTACCCCA	1200
Db	3068	AATGGCTGCCCGCCACCCCGGGCTCCCTTGGGCAAAAGAAATGTCAGCCCTACCCCA	3127
QY	1201	ACCCTTCAACTACAGATCTGGGCCACCCAGAGATTTTATTTAAATGTTGCCCA	1260
Db	3128	ACCCTTCAACTACAGATCTGGGCCACCCAGAGATTTTATTTAAATGTTGCCCA	3187
QY	1261	TTTTATGAGTTATGATCAATTTGATTTAAATTTAAAGTTACAGATGTCAAAAA	1320
Db	3188	TTTTATGAGTTATGATCAATTTGATTTAAATTTAAAGTTACAGATGTCAAAAA	3247
QY	1321	AAAAAAAA 1327	
Db	3248	AAAAAAAA 3254	
RESULT 5			
AX491279	LOCUS	AX491279	DNA linear PAT 16-AUG-2002
DEFINITION	Sequence 3 from Patent WO0212285.		
ACCESSION	AX491279		
VERSION	AX491279.1 GI:22323980		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Killary, A., Chandler, D. and Lott, S.		
TITLE	The tumor suppressor car-1		
JOURNAL	Patent: WO 0212285-A 3 14-FEB-2002;		
FEATURES	BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)		
Source	Location/Qualifiers		
	1..3826		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	100.0%; Score 1327; DB 6; Length 3826;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches 1327; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	TCACAGTGTGCAGGTAGTACTGTTCTAGGGTTGCTTGAGAGCCACCTCTCTGC	60

REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS	1 (bases 1 to 164950) Waterston,R.H.
TITLE	The sequence of Homo sapiens clone
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 164950) Waterston,R.H.
AUTHORS	Direct Submission
TITLE	Submitted (27-JAN-2000) Genome Sequencing Center, Washington
JOURNAL	University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
COMMENT	On Jun 16 2000 this sequence version replaced gi:73219234.


```
Db 17561 TCTCTGACACACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGTTCTTAA 17502
Qy 1141 AATGGCTGCCCGCCACCGGGCTCCCTTTGGGCAAAAGAAATGTCAGCCCTACCCCA 1200
Db 17501 AATGCTGCCCGCCACCGGGGCTCCCTTTGGGCAAAAGAAATGTCAGCCCTACCCCA 17442
Qy 1201 ACCCTTCAACTACAGAACTCTGGGCCACCCCGAGCAGTATTTTATTTAAATGTTGCCA 1260
Db 17441 ACCCTTCAACTACAGAACTCTGGGCCACCCCGAGCAGTATTTTATTTAAATGTTGCCA 17382
Qy 1261 TTTTATGAGTTATGATCAATTTGTTATTTAAATTAAGTTTACAGATGTC 1308
Db 17381 TTTTATGAGTTATGATCAATTTGTTATTTAAATTAAGTTTACAGATGTC 17334

RESULT 8
AX775831
LOCUS AX775831 3243 bp mRNA linear PAT 14-JUL-2003
DEFINITION Sequence 101 from Patent WO03048202.
ACCESSION AX775831
VERSION AX775831.1 GI:32693549
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Matsuda,A. and Muramatsu,S.
TITLE NF-kB activating gene
JOURNAL Patent: WO 03048202-A 101 12-JUN-2003;
Asahi Kasei Kabushiki Kaisha (JP)
FEATURES
Location/Qualifiers
1..3243
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
43..1470
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAB11615.1"
/db_xref="GI:32693550"
/translation="MACSKDELKLSICLSIYQDPVSLGCEHYFCRRCTTEHWVROEA
QGARDCPECRTPAEALAPSLKLANIVERYSSFPPLDALINARRAARPCQAHDKVLF
CLTDRLALCFCDPEALHEQOVTGIDDAFDELQRELKDQLOLQDSEHEHEALQIL
KCOLATKSKTSKSLRTTIGEARLRLRLREROKAMLEBDAFLRLDIDIEOKVRY
SQOLRVQGAQILQRLAETDRHFLAGVASLSERLKGKIHETNLITYEDFPSTKYG
PIQYTIWKLFLQDHPVPAALTLDPGTARLQRLISDDCTIVAYGNLHPQLDPSKRF
DVEVSLGSEAFSSGVHYVEVVAEKTOQWVIGLAHEAASRKGSIQIQRGRFYCIWMH
DGNQYACTEPWTRLNVRDKLVGVFLDQGLLIFYNADDMSLWLYTFREKFPGLC
SVFSPQSHANGKNVQPLRINTVRI"
ORIGIN
Query Match 98.5%; Score 1307; DB 6; Length 3243;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TCACAGTGTGCACAGGTAGTACTGGTCTCTAGGGTTGGCTGAGAGCCAACTCTCTCTGC 60
Db 1937 TCACAGTGTGCACAGGTAGTACTGGTCTCTAGGGTTGGCTGAGAGCCAACTCTCTCTGC 1996
Qy 61 CACCCCCACACCAAGAACTATATGTTTCTTCTCTCCCACTGATCTGCTGGTTCAGTGAT 120
Db 1997 CACCCCCACACCAAGAACTATATGTTTCTTCTCTCCCACTGATCTGCTGGTTCAGTGAT 2056
Qy 121 GATGCTGTGGCTGTGGAGGCAACCTGGTAGTTGAGTCCACACATATATAGTCATGTGCCA 180
Db 2057 GATGCTGTGGCTGTGGAGGCAACCTGGTAGTTGAGTCCACACATATATAGTCATGTGCCA 2116
Qy 181 CCACCTTCTCTGCCACAGCCGAGGACAGGGTGAAGGTATACCCAAAGCTGATGCAGAG 240
Db 2117 CCACCTTCTCTGCCACAGCCGAGGACAGGGTGAAGGTATACCCAAAGCTGATGCAGAG 2176
```

```
Qy 241 CCCATTAGCTTAAAGCAACTGCAAGCAACGCTTCCCTGTGATGATCGAGGTCCCCAGTAG 300
Db 2177 CCCATTAGCTTAAAGCAACTGCAAGCAACGCTTCCCTGTGATGATCGAGGTCCCCAGTAG 2236
Qy 301 CTCTGAACAAGAGTCCAGCAACCCCTCTTCAAGCAGAGCTCTGTGACTGTGAGGTGCA 360
Db 2237 CTCTGAACAAGAGTCCAGCAACCCCTCTTCAAGCAGAGCTCTGTGACTGTGAGGTGCA 2296
Qy 361 GGAGGCTTCCAGAGCAGTGTGTTGTAATTAGGACCCCAAGCACTGGGAGGGGCTGTGGCT 420
Db 2297 GGAGGCTTCCAGAGCAGTGTGTTGTAATTAGGACCCCAAGCACTGGGAGGGGCTGTGGCT 2356
Qy 421 AGACCCCTGTGTCAGACTTGCATCTATCTCAGTTAGGATCTCTGTGAGAGAAAACAAGAGC 480
Db 2357 AGACCCCTGTGTCAGACTTGCATCTATCTCAGTTAGGATCTCTGTGAGAGAAAACAAGAGC 2416
Qy 481 CACTTGTAGCTGGTGTAAATTAGACAAGGATTTACTACTTGCCTGSCCTCTGGTGTGCAAAA 540
Db 2417 CACTTGTAGCTGGTGTAAATTAGACAAGGATTTACTTCTGCCCCCTGGTGGCTTGCAAA 2476
Qy 541 TTGTTGAAGAGCTGGAGAGCAGACTCTGCTGAAATTTCCAGGAACCTCCAGGCGCCAGAT 600
Db 2477 TTGTTGAAGAGCTGGAGAGCAGACTCTGCTGAAATTTCCAGGAACCTCCAGGCGCCAGAT 2536
Qy 601 TCATCATGCTGTGTGACCAAGAAAGCTGCCCCCATCTCTGCAGGAAGCCACTATGCGCAGA 660
Db 2537 TCATCATGCTGTGTGACCAAGAAAGCTGCCCCCATCTCTGCAGGAAGCCACTATGCGCAGA 2596
Qy 661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCGCCAGGTCGCTGCGCAGCCCAATAGATGTC 720
Db 2597 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCGCCAGGTCGCTGCGCAGCCCAATAGATGTC 2656
Qy 721 CTGAGGCTGCCCCCTCTCCCACTTCACTCAGTTTCCCAAAATCTAAATTTTACAAAGATTT 780
Db 2657 CTGAGGCTGCCCCCTCTCCCACTTCACTCAGTTTCCCAAAATCTAAATTTTACAAAGATTT 2716
Qy 781 CTGTTTGGGGGAACCTTAAGTCAGATCCAGAACTTGGCTGCAAGGGAGTCTGGAAATGT 840
Db 2717 CTGTTTGGGGGAACCTTAAGTCAGATCCAGAACTTGGCTGCAAGGGAGTCTGGAAATGT 2776
Qy 841 CATTTCCTTAGAAGGAAGTTAGGCTGGTGGAGCAAGCCCACTGGTGTCTTCTGCCAC 900
Db 2777 CATTTCCTTAGAAGGAAGTTAGGCTGGTGGAGCAAGCCCACTGGTGTCTTCTGCCAC 2836
Qy 901 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCTTGCCC 960
Db 2837 AGCATCCAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATCTAGGGTTGTCTTGCCC 2896
Qy 961 CTTGGCTCTATCCCTGCCAGAGTGGGAACTGGAGAGTGGGCTGCAAGACTGAGCCTA 1020
Db 2897 CTTGGCTCTATCCCTGCCAGAGTGGGAACTGGAGAGTGGGCTGCAAGACTGAGCCTA 2956
Qy 1021 AATGCTCTCCCGGCTTGTGACTTTTCTTCTTAGTTCCTGGGCGCTAGATTTCTGCATTTGGGG 1080
Db 2957 AATGCTCTCCCGGCTTGTGACTTTTCTTCTTAGTTCCTGGGCGCTAGATTTCTGCATTTGGGG 3016
Qy 1081 TCTCTGACACAACACACCACTCCCAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCCTTAA 1140
Db 3017 TCTCTGACACAACACACCACTCCCAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCCTTAA 3076
Qy 1141 AATGCTCTCCCGGCTGCAAGAGTGGGCTCTTGGGCAAAAGAAATGTACGCCCTACCCCA 1200
Db 3077 AATGCTCTCCCGGCTGCAAGAGTGGGCTCTTGGGCAAAAGAAATGTACGCCCTACCCCA 3136
Qy 1201 ACCCTTCAACTACCAAGAACTCTGGGCCACCCCGAGCAGTATTTTATTTAAATGTTGCCA 1260
Db 3137 ACCCTTCAACTACCAAGAACTCTGGGCCACCCCGAGCAGTATTTTATTTAAATGTTGCCA 3196
Qy 1261 TTTTATGAGTTATGATCAATTTGTTATTTAAATTAAGTTTACAGATGTC 1307
Db 3197 TTTTATGAGTTATGATCAATTTGTTATTTAAATTAAGTTTACAGATGTC 3243
```



```

RESULT 9
AX491284
LOCUS      AX491284
DEFINITION Sequence 8 from Patent WO0212285.
ACCESSION AX491284
VERSION    AX491284.1 GI:22323985
KEYWORDS   Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
REFERENCE   1
AUTHORS     Killary, A., Chandler, D. and Lott, S.
TITLE       The tumor suppressor car-1
JOURNAL     Patent: WO 0212285-A B 14-FEB-2002;
            BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)
FEATURES    Location/Qualifiers
             1..30676
                /organism="Homo sapiens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:9606"

ORIGIN
Query Match      94.7%; Score 1257; DB 6; Length 30676;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1307; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TCCACAGTGTGTCACAGGTAGTACCTGCTTCTAGGTTGCTGAGGCGCAACCTCTCTGCG 60
Db 24440 TCCACAGTGTGTCACAGGTAGTACCTGCTTCTAGGTTGCTGAGGCGCAACCTCTCTGCG 24499

Qy 61 CACCCGCCACACAGACTATATGTTCTCTACTTCTCCACTGATCTGCTGCTGCTGAT 120
Db 24500 CACCCGCCACACAGACTATATGTTCTCTACTTCTCCACTGATCTGCTGCTGCTGAT 24559

Qy 121 GATGCTGTGGCTGTGGAGGCACTGCTGTAGTTCAGTCCACACATTATAGTCATGTGCCA 180
Db 24560 GATGCTGTGGCTGTGGAGGCACTGCTGTAGTTCAGTTCACACATTATAGTCATGTGCCA 24619

Qy 181 CCACCTTCTCTCCACAGCCGAGGCAAGGCTGAGGTATACCCAAAGCTGATGCGAG 240
Db 24620 CCACCTTCTCTCCACAGCCGAGGCAAGGCTGAGGTATACCCAAAGCTGATGCGAG 24679

Qy 241 CCCATTAGCTTAAGCAAGCACTGCAGGCAAGCCTCCCTGGATGATCGAGGTCCTCCAGTAG 300
Db 24680 CCCATTAGCTTAAGCAAGCACTGCAGGCAAGCCTCCCTGGATGATCGAGGTCCTCCAGTAG 24739

Qy 301 CTCTGAAACAAGAGTCCAGCCCAACCTCTTTAGCCAGGCTCTGTGACCTGTAGGCTGCA 360
Db 24740 CTCTGAAACAAGAGTCCAGCCCAACCTCTTTAGCCAGGCTCTGTGACCTGTAGGCTGCA 24799

Qy 361 GGAGGCTTCCAGAGCAGTGTGTGTAATAGAGCCCAAGCACTGGAGGGGCTGTGGCT 420
Db 24800 GGAGGCTTCCAGAGCAGTGTGTGTAATAGAGCCCAAGCACTGGAGGGGCTGTGGCT 24859

Qy 421 AGACCCCTGTGACAGCTGGCATCTATCTCAGTATAGGATCTGCTGCGCAAAACAGAGC 480
Db 24860 AGACCCCTGTGACAGCTGGCATCTATCTCAGTATAGGATCTGCTGCGCAAAACAGAGC 24919

Qy 481 CACTTGTAGCTGGTGTAAATAGACAGGATTTACTACCTGGCCCTCTGGTGGCTTGC AAA 540
Db 24920 CACTTGTAGCTGGTGTAAATAGACAGGATTTACTACCTGGCCCTCTGGTGGCTTGC AAA 24979

Qy 541 TTGTTGGAGAGCTGGAGAGCAGACTCTGCTGAATTTCCAGGAATCTCCAGGCGCAGAT 600
Db 24980 TTGTTGGAGAGCTGGAGAGCAGACTCTGCTGAATTTCCAGGAATCTCCAGGCGCAGAT 25039

Qy 601 TCATCATGCTGTGTGACAGAGAGCTGCCCCCATCTGACAGGAGGCACTATGCGAGA 660
Db 25040 TCATCATGCTGTGTGACAGAGAGCTGCCCCCATCTGACAGGAGGCACTATGCGAGA 25099

Qy 661 AAGCTGCTGACTGCAAGCTAGGCTCCCTCTGCGCAGGTCCTGCGCAGGCAATAGATGTC 720

```

```

Db 25100 AAGCTGCTGACTGCAAGAACTAGGCTCCCTCTGCCACGGTCCGCTGCCAGCAATAGATGTC 25159
Qy 721 CTGAGGCGCTGCCCTCTCCCACTTCACTCAGTTTCCCAAAATCTAAATTTTACAGAGATT 780
Db 25160 CTGAGGCGCTGCCCTCTCCCACTTCACTCAGTTTCCCAAAATCTAAATTTTACAGAGATT 25219
Qy 781 CTGTTTGGGGGAACTTAAAGTCAGATCCAGAACCTTGCTGCTCAAGGGAGTCTGGAAATGT 840
Db 25220 CTGTTTGGGGGAACTTAAAGTCAGATCCAGAACCTTGCTGCTCAAGGGAGTCTGGAAATGT 25279
Qy 841 CATTTCCCTAGAGGAAGTTAGGGTGGAGGAGAGCCCACTGGCTTTTCTGCGCAC 900
Db 25280 CATTTCCCTAGAGGAAGTTAGGGTGGAGGAGAGCCCACTGGCTTTTCTGCGCAC 25339
Qy 901 AGCATCCAATCGTGAAGAACTCCGGAGAGGTTGGAGTCCACATCTAGGTTGTCTGCTGCC 960
Db 25340 AGCATCCAATCGTGAAGAACTCCGGAGAGGTTGGAGTCCACATCTAGGTTGTCTGCTGCC 25399
Qy 961 CTTGGCTCTATCCCTGCCAGAGTGGAACTGGAGGAGTGGGCTGCAAGACTGAGGCTA 1020
Db 25400 CTTGGCTCTATCCCTGCCAGAGTGGAACTGGAGGAGTGGGCTGCAAGACTGAGGCTA 25459
Qy 1021 AATGCTCCTCCCGGCTTGACTTTTCTTTCTAGTCTGGGCTAGATTTCTGCACTGGGG 1080
Db 25460 AATGCTCCTCCCGGCTTGACTTTTCTTTCTAGTCTGGGCTAGATTTCTGCACTGGGG 25519
Qy 1081 TCTCTGACACAAACACACATCCCAAGTAGCCGGAAGAGCTTAAACACAGGGGGTCTTTAA 1140
Db 25520 TCTCTGACACAAACACACATCCCAAGTAGCCGGAAGAGCTTAAACACAGGGGGTCTTTAA 25579
Qy 1141 AATGGCTGCCCGCCCGCCAGGCTCCCTTTGGGCAAAAGAAATGTCAGCCCTACCCCA 1200
Db 25580 AATGGCTGCCCGCCCGCCAGGCTCCCTTTGGGCAAAAGAAATGTCAGCCCTACCCCA 25639
Qy 1201 ACCCTTCAACTACAGAACTCTGGGCCACCCAGCAGTATTTTATTAATGTTGCCCCA 1260
Db 25640 ACCCTTCAACTACAGAACTCTGGGCCACCCAGCAGTATTTTATTAATGTTGCCCCA 25699
Qy 1261 TTTTATGAGTTATGATCAATTTGTTAAATTTAAAGTTACAGATGTC 1308
Db 25700 TTTTATGAGTTATGATCAATTTGTTAAATTTAAAGTTACAGATGTC 25747

```

```

RESULT 10
AL662907/c
LOCUS      AL662907
DEFINITION Human DNA sequence from clone RP11-131M11 on chromosome 1, complete
            sequence.
ACCESSION AL662907 AC026053
VERSION    AL662907.11 GI:21531085
KEYWORDS   HTG
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 64693)
AUTHORS     Van Hellmond, Z.
TITLE       Direct Submission
JOURNAL     Submitted (18-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquerry@sanger.ac.uk
            On Jun 21, 2002 this sequence version replaced gi:20145349.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least

```

one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/c_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1> RP11-131M11 is from the library RP11-11.1 constructed by the group of Peter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBAC3.6 Draft Sequence Produced by Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA <http://genome.wustl.edu/gsc/index.shtml>.

FEATURES

source
1..64693
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-131M11"
/clone_lib="RP11-11.1"

ORIGIN

Query Match 94.7%; Score 1257; DB 9; Length 64693;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1307; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1 TCACACGTGGTTCACAGGTAGTACCTGCTCTCTAGGTTGGCTGAGAGCAACCTCTCTGC 60
Db TCACAGTGGTTCACAGGTAGTACCTGCTCTAGGTTGGCTGAGAGCAACCTCTCTGC 7173
61 CACCCACACACCAAGAACTATATGTTCTCTACTTCTCCACTGATCTGCTGGTCAGTGAT 120
Db CACCCACACACCAAGAACTATATGTTCTCTACTTCTCCACTGATCTGCTGGTCAGTGAT 7113
121 GATGCTGTGGCTGTGGAAGGACCTGGTGTGAGTCCACATATATGTCATGTGCCA 180
Db GATGCTGTGGCTGTGGAAGGACCTGGTGTGAGTCCACATATATGTCATGTGCCA 7053
181 CCACCTTCTCTGCCACAGGCGGAGGACAGGTGAGGTATACCAAGCTGTGACAG 240
Db CCACCTTCTCTGCCACAGGCGGAGGACAGGTGAGGTATACCAAGCTGTGACAG 6993
241 CCATTAGCTTAAAGCAACTGCAGGACAAAGCTCTCCCTGGATGATCGAGTCCCCAGTAG 300
Db CCATTAGCTTAAAGCAACTGCAGGACAAAGCTCTCCCTGGATGATCGAGTCCCCAGTAG 6933
301 CTCTGACAGAGTCCAGCAACCTCTTCCAGGAGGCTCTGTGACCTCTAGGGTGCA 360
Db CTCTGACAGAGTCCAGCAACCTCTTCCAGGAGGCTCTGTGACCTCTAGGGTGCA 6873
361 GGAGGCTTCCAGAGCAGTGTGTTGTAATTAGGACCCAGCACTGGGAGGGGCTGTGGCT 420
Db GGAGGCTTCCAGAGCAGTGTGTTGTAATTAGGACCCAGCACTGGGAGGGGCTGTGGCT 6813
421 AGACCCCTTGTACAGCTTGCATCTATCTCAGTTAGGATCCTCTGCAGAAACCAAGAGC 480
Db AGACCCCTTGTACAGCTTGCATCTATCTCAGTTAGGATCCTCTGCAGAAACCAAGAGC 6753
481 CACTTGTAGCTGTTTAAATAGACAGGATTTACTACCTGGCCCTGGTGGCTGCAAAA 540
Db CACTTGTAGCTGTTTAAATAGACAGGATTTACTACCTGGCCCTGGTGGCTGCAAAA 6693
541 TTCTTGAAGAGCTGAGAGCAGACTCTGCTGAATTTCCAGAACTCCAGGCCAGAT 600
Db TTCTTGAAGAGCTGAGAGCAGACTCTGCTGAATTTCCAGAACTCCAGGCCAGAT 6633

QY 601 TCATCATGCTCTGTTGTGACCAAGAAAGCTGCCCATCTGCGAGGAGCACTATGCCAGA 660
Db 6632 TCATCATGCTCTGTTGTGACCAAGAAAGCTGCCCATCTGCGAGGAGCACTATGCCAGA 6573
QY 661 AAGTCTGCTGACTGCAGAACTAGGCTCCCTCTGCGACGCTCCGTCGCCCAATAGATGTC 720
Db 6572 AAGTCTGCTGACTGCAGAACTAGGCTCCCTCTGCGACGCTCCGTCGCCCAATAGATGTC 6513
QY 721 CTGAGGGCTGCCCTCTCCCACTTCACTCACTTCCCAATCTAAATTTTACAAGAGATT 780
Db 6512 CTGAGGGCTGCCCTCTCCCACTTCACTCACTTCCCAATCTAAATTTTACAAGAGATT 6453
QY 781 CTGTTTGGGGAACCTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
Db 6452 CTGTTTGGGGAACCTTAAGTCAGATCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 6393
QY 841 CATTTCCCTGAAGAAAGTTAGGCTGGTGGAGCAAGCCACCTGGTCTTTCTGCCAC 900
Db 6392 CATTTCCCTGAAGAAAGTTAGGCTGGTGGAGCAAGCCACCTGGTCTTTCTGCCAC 6333
QY 901 AGCATCCAAATCGTGAAGAACTCGGAGAGGGTGGAGTCCACATCTAGGGTTGTCTGCC 960
Db 6332 AGCATCCAAATCGTGAAGAACTCGGAGAGGGTGGAGTCCACATCTAGGGTTGTCTGCC 6273
QY 961 CTTGGCTCTATCCCTGCCAGAGTGGAACTGAGGAGTGGGTGCAAGCTAGGCTA 1020
Db 6272 CTTGGCTCTATCCCTGCCAGAGTGGAACTGAGGAGTGGGTGCAAGCTAGGCTA 6213
QY 1021 AATGTCTCCCGGGCTTGACTTTTCTAGTCTCTGGGCTAGATTCTGCATTTGGG 1080
Db 6212 AATGTCTCCCGGGCTTGACTTTTCTAGTCTCTGGGCTAGATTCTGCATTTGGG 6153
QY 1081 TCTCTGACACAAACACACCATCCCAAGTAGCCGAGAGCTAAACACAGGGGTTCTTAA 1140
Db 6152 TCTCTGACACAAACACACCATCCCAAGTAGCCGAGAGCTAAACACAGGGGTTCTTAA 6093
QY 1141 AATGGCTGCCCGCCACACCGGGCTCCCTTGGGCAAAAAGAAATGTGACCCCTACCCCA 1200
Db 6092 AATGGCTGCCCGCCACACCGGGCTCCCTTGGGCAAAAAGAAATGTGACCCCTACCCCA 6033
QY 1201 ACCCTTCAACTACAGAACTCTGGGCGACCCCGAGCAGTATTTTATTAAGTGTGCCCA 1260
Db 6032 ACCCTTCAACTACAGAACTCTGGGCGACCCCGAGCAGTATTTTATTAAGTGTGCCCA 5973
QY 1261 TTTTATGAGTATGATCAATTTCTATTAATAAAGTTACAGATGCA 1308
Db 5972 TTTTATGAGTATGATCAATTTCTATTAATAAAGTTACAGATGCA 5925

RESULT 11

AX877131
LOCUS AX877131
DEFINITION Sequence 12036 from Patent BP1074617.
ACCESSION AX877131
VERSION AX877131.1 GI:40031867
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ota.T., Isogai.T., Nishikawa.T., Hayashi.K., Saito.K., Yamamoto.J., Ishii.S., Sugiyama.T., Wakamatsu.A., Nagai.K. and Otsuki.T.
TITLE Primers for synthesising full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 12036 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
source
1..3436
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
235..1662
/note="unnamed protein product"

```

/codon_start=1
/protein_id="CAB89585.1"
/db_xref="GI:40031868"
/translation="MACS3KDELLCSICLSTYQDPVSLGCEHYFCRRCTEHWROEA
QAGARQCEPRFASFAPALPSKLIANIVYSSPFLDIALNARARPCQADHKVLL
CTDUALLCFCDEPAAHQHVQTVIDAFDELQKLELDQALQASREHTALQLL
KQLATLTKSSTKSLATLTGGEAFRLHRLRROKAMELEETADTALTIDEQVRY
SQQLKVBQEGAIQIERLAETDRTFTAGVASLRLKGIHNTLVYDFPQSKYTG
PQYTLWKSLFQDHPFAAUGLDPGTAHQRLILNDSCTVAGNHLHPQLQDSKPRF
DVEVYLGSFAEPVWVAAKCTQWVIGLHAASRKGSIQPSRQSCFVCIVMH
DNGQKSACTEPTKINVRKLDKGVDFDLYDQGLLITFNADDSWLVYFREKFPGLC
SVFPGSGAGHNKVPQIRLNTVL"
```

ORIGIN

Query Match	94.6%	Score 1256	DB 6	Length 3436
Best Local Similarity	99.9%	Pred. No. 0		
Matches 1306	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	1	TCCACAGTGGTACACAGTAGTACCTGGTCTAGGTTGCTCAGAGCCAACTCTCTGC	60	
Db	2130	TCCACAGTGGTACACAGTAGTACCTGGTCTAGGTTGCTCAGAGCCAACTCTCTGC	2189	
QY	61	CAGCCCCACACAGAACTATATGGTTCTACTCTCCACATGATCTGCTGTCAGTGAT	120	
Db	2190	CAGCCCCACACCAAGAACTATATGGTTCTACTCTCCACATGATCTGCTGTCAGTGAT	2249	
QY	121	GATGCTGTGGCTGTGGAGGACACCTGGTAGTTGAGTCCACACATATATAGTATGCGCA	180	
Db	2250	GATGCTGTGGCTGTGGAGGACACCTGGTAGTTGAGTCCACACATATATAGTATGCGCA	2309	
QY	181	CCACCTTCTCTGCCACACAGCCGAGGACAGGGTGAGGGTATACCCAAAGTGATGACAG	240	
Db	2310	CCACCTTCTCTGCCACACAGCCGAGGACAGGGTGAGGGTATACCCAAAGTGATGACAG	2369	
QY	241	CCCATTAGCTTAAGACCACTGCAGACACAGCCCTCCCTGGATGATCAGAGTCCCAAGTAG	300	
Db	2370	CCCATTAGCTTAAGACCACTGCAGACACAGCCCTCCCTGGATGATCAGAGTCCCAAGTAG	2429	
QY	301	CTCTGAACAGAGTCCAGGCCAACCCCTCTTCAGCCAGGCGCTCTGTGACCTGCTAGGTGCA	360	
Db	2430	CTCTGAACAGAGTCCAGGCCAACCCCTCTTCAGCCAGGCGCTCTGTGACCTGCTAGGTGCA	2489	
QY	361	GGAGGCTTCAGAAACAGTGTGTGTAATTAGGACCCAGCACTGGGAGGGCTGTGGCT	420	
Db	2490	GGAGGCTTCAGAAACAGTGTGTGTAATTAGGACCCAGCACTGGGAGGGCTGTGGCT	2549	
QY	421	AGACCCCTTCTCAGACTTGGCATCTATCTCAGATTAGGATCCTGCTGCAGAAAAACAAGAC	480	
Db	2550	AGACCCCTTCTCAGACTTGGCATCTATCTCAGATTAGGATCCTGCTGCAGAAAAACAAGAC	2609	
QY	481	CACCTGAGCTGGTTTAATTAGACAGGATTTACTACTGGCCCTCTGGTGGCTTGCAAAA	540	
Db	2610	CACCTGAGCTGGTTTAATTAGACAGGATTTACTACTGGCCCTCTGGTGGCTTGCAAAA	2669	
QY	541	TTCTGTGAAGAGCTGGAGAGCAGACTCTCTGTAATTTCCAGGAATCCCCAGCGCCAGAT	600	
Db	2670	TTGTGTGAAGAGCTGGAGAGCAGACTCTCTGTAATTTCCAGGAATCCCCAGCGCCAGAT	2729	
QY	601	TCATCATGCTGTGTGTGACCAAGAAAGCTGCCCCCATCTGCAGAAAGCCACTATGCCAGA	660	
Db	2730	TCATCATGCTGTGTGTGACCAAGAAAGCTGCCCCCATCTGCAGAAAGCCACTATGCCAGA	2789	
QY	661	AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCAATAGATGTC	720	
Db	2790	AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCAATAGATGTC	2849	
QY	721	CTGAGGCTGTGCCCTCTCCCACTTCACTCAGTTCCCAAACTCTAAATTTTACAGAGATT	780	
Db	2850	CTGAGGCTGTGCCCTCTCCCACTTCACTCAGTTCCCAAACTCTAAATTTTACAGAGATT	2909	
QY	781	CTGTTTGGGGAACTTAAGTCAGATCCAGAAACCTTGGCTGCAAGGAGCTCGGGAAATGT	840	
Db	2910	CTGTTTGGGGAACTTAAGTCAGATCCAGAAACCTTGGCTGCAAGGAGCTCGGGAAATGT	2969	

Qy	841	CA	TTTCCCTAGAAAGGAAGTATGGGTGGGTGGAGCAAGCCCACTCGTGGTTTTCTGCCAC	900
Db	2970	CA	TCTCCCTTAGAAAGGAAGTATGGGTGGGTGGAGCAAGCCCACTCGTGGTTTTCTGCCAC	3029
Qy	901	AG	CATCCAAATCGTGAAGAACTCGGAGAGAGGGTGGAGTCCACATCTAGGGTTGCTTGCCC	960
Db	3030	AG	CATCCAAATCGTGAAGAACTCGGAGAGAGGGTGGAGTCCACATCTAGGGTTGCTTGCCC	3089
Qy	961	CT	TGGCTCTATCCCTCCAGAGGTGGGAACCTGGAGAGTGGGCTGCAAGACTGAGCCTA	1020
Db	3090	CT	TGGCTCTATCCCTCCAGAGGTGGGAACCTGGAGAGTGGGCTGCAAGACTGAGCCTA	3149
Qy	1021	AA	TGCTCCCGGCTTGACTTTCTTTCTAGTCCTGGGGCCCTAGATTCTGCACCTTGGGG	1080
Db	3150	AA	TGCTCCCGGCTTGACTTTCTTTCTAGTCCTGGGGCCCTAGATTCTGCACCTTGGGG	3209
Qy	1081	TC	TCTGACACAACACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCCTAA	1140
Db	3210	TC	TCTGACACAACACACACCATCCCAAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCCTAA	3269
Qy	1141	AA	TGGTGTGCCCCGACACCGGGCCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA	1200
Db	3270	AA	TGGTGTGCCCCGACACCGGGCCCTCCCTTGGGCAAAAGGAATTGTCAGCCCTACCCCA	3329
Qy	1201	AC	CCCTTCAACTACCAAGATCTGGGGCAACCCAGCAGTATTTTATTTAAATGTTGCCCA	1260
Db	3330	AC	CCCTTCAACTACCAAGATCTGGGGCAACCCAGCAGTATTTTATTTAAATGTTGCCCA	3389
Qy	1261	TT	TATGAGTTATGATCAATTCGTATTAATTAAGTTACAGATGTC	1307
Db	3390	TT	TATGAGTTATGATCAATTCGTATTAATTAAGTTACAGATGTC	3436

REFUGEE 12

BD156501	3436 bp	DNA	linear	PAT 17-JAN-2003
Primer for synthesizing full-length cDNA and use thereof.				
BD156501				
BD156501.1	GI:27862259			
JP 2002191363-A/11344.				
Homo sapiens (human)				
Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
1 (bases 1 to 3436)				
Ota,T., Isegai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.				
Primer for synthesizing full-length cDNA and use thereof				
Patent: JP 2002191363-A 11344 09-JUL-2002;				
HELIX RESEARCH INSTITUTE				
OS Homo sapiens (human)				
FN JP 2002191363-A/11344				
PD 09-JUL-2002				
PF 28-JUL-2000 JP 2000280990				
PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAOJU SAITO.				
PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU, PI KETICHI NAGAI,TETSUJI OTSUKI				
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC 10,				
PC, C12P21/02,C12O1/68/C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC				
Primer for synthesizing full-length cDNA and use thereof FH Key				
Location/Qualifiers				
(235)..(1659).				
FT CDS				
Location/Qualifiers				
1..3436				
/organism="Homo sapiens"				
/mol_type="genomic DNA"				
/db_xref="taxon:9606"				
FEATURES				
source				
ORIGIN				

ORIGIN

Query Match	94.6%;	Score 1256;	DB 6;	Length 3436;
Best Local Similarity	99.9%;	Prod. No. 0;		
Matches 1306;	Conservative	0;	Mismatches 1;	Indels 0;
				Gaps 0;
QY	1	TCCACAGTGGTTCACAGGTAGTACCTGGTCTTAGGGTTGCCTGAGAGCCAACTCTCTCTGC	60	
DB	2130			
QY	61	CACCCCCACACAAAGAACTATATGGTTTCTACTTCTCCCACTGATCTCTGCTGCTCAGTGAT	120	
DB	2190			
QY	121	GATGCTGTGGCTGTGGGAAGCACTGTGTAGTTGAGTCCACACATTTAGTCATGTGCGAT	2249	
DB	2250			
QY	181	CCACTTCTCCGCCACAGGCCGAGGACAGGTCAGGTGAGGTATACCCAAAGCTGATGCGCA	180	
DB	2310			
QY	241	CCCATTAGCCTTAAAGCAACTGCAAGCAACAAGCCTCCCTGGAGTATGTCAGAGTCCCCACTAG	300	
DB	2370			
QY	301	CTCTGAACAAGAGTCCAGCCAAACCTTTTCAGCCAGGCTCTGTGACCTGCTAGGGTGCA	360	
DB	2430			
QY	361	GGAGGCTTCAGAAGACGAGTTGTCGTATATAGGACCCAAAGCACTGGGAGGGGCTGTTGGCT	420	
DB	2490			
QY	421	AGACCCCTTGTTCAGACTTGGCACTATATCTCAGTTTAGGATCTGCTGTCGAGAAAAACAAGAC	480	
DB	2550			
QY	481	CACTTGTAGCTGGTTTAAATTAGACAAGGATTTACTCTGGCCCCCTGGTGGCTTGCAAAA	540	
DB	2610			
QY	541	TTGTTGGAAGAGCTGGGAAGACAGACTCTGCTGAAATTTCCAGGAACTCCCGAGCGCCAGAT	600	
DB	2670			
QY	601	TCATCATGTCTGTTGTGACCAGGAAAGCTGCCCCCATCTGCAGGAAGCCACTATGCCAGA	660	
DB	2730			
QY	661	AAGTCGTGACTGCGAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCCAAATAGATGTC	720	
DB	2790			
QY	721	CTGAGGCTCCCCCTCTCCCACTTCACTCAGTTCCCAATCTAAAATTTTACAAGAGATT	780	
DB	2850			
QY	781	CTGTTTGGGGAACTTAAAGTCAGATCCAGAACTTTGGCTCAAGGGAGTCTGGGAATGT	840	
DB	2910			
QY	841	CATTTCCCTAGAGGAAGTTAGGTTGGGTGGAGCAAGCCCCCACTCGTTTTTCTGCCAC	900	
DB	2970			
QY	901	AGCATCCMATCGTGAAGAACTCGGAGAGGGTGGAGTCCACATCTTAGGTTGTCTGCC	960	
DB	3030			
QY	961	CTTGGCTCTATCCCTGCCCCAGAGTGGAACTGGAGGAGTGGGCTCAGACTGAGCCTA	1020	
DB	3090			
QY	1021	AATGTCCTCCCGCCTTGACTTTTCTTCTTAGTCTCTGGGCGCTTAGATTCTGCACTTGGGG	1080	

ACCESSION	AX491281	23851	AGCCCCACCTGCGTTTTTTCTGCCACAGCATCCCAATCGTGAAAAAATCTCGGAGAGGGGTGGA	23911
VERSION	AX491281.1	936	GTCCACATCTAGGGTTGTCTGCCCCCTTGCTCTATCCCTGCCAGAGGTGGGAACCTGGA	995
KEYWORDS		23911	GTCCACATCTAGGGTTGTCTGCCCCCTTGCTCTATCCCTGCCAGAGGTGGGAACCTGGA	23970
SOURCE	Homo sapiens (human)	996	GGAGTGGGCTGCAAGACTGAGCCCTAAATGCTCCCGGCTTGACTTTCTTTCTAGTCC	1055
ORGANISM	Homo sapiens	23971	GGAGTGGGCTGCAAAACTGAGCCCTAAATGCTCCCGGCTTGACTTTCTTTCTAGTCC	24030
REFERENCE	Killary,A., Chandler,D. and Lott,S. The tumor suppressor car-1 Patent: WO 021285-A 5 14-FEB-2002; BOARD OF REGENTS THE UNIVERSITY OF TEXAS SYSTEM (US)	1056	TGGGGCCCTAGATCTGCACTTGGGCTCTCTGACACACACACACATCCCAAAGTAGCCGGA	1115
AUTHORS		24031	TGGGGCCCTAGATCTGCACTTGGGCTCTCTGACACACACACACATCCCAAAGTAGCCGGA	24090
TITLE		1116	AGAGCTAAACACACAGGGGTTCTTAAATGGTGTGCCCCGCCACCGGCTCCCTTGGGC	1175
JOURNAL		24091	AAAGCTAAACACACAGGGGTTCTTAAATGGTGTGCCCCGCCACCGGCTCCCTTGGGC	24150
FEATURES	Location/Qualifiers 1..30625 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"	1176	AAAAGGAATTGTGAGCCCTACCCCAACCCCTTCAACTACCAATCTGGGCCACCCGACGA	1235
source		24151	AAAAGGAATTGTGAGCCCTACCCCAACCCCTTCAACTACCAATCTGGGCCACCCGACGA	24210
ORIGIN		1236	GTATTTTATTAATAATGTTGCCCATTTTATGAGTTATGATCAATTTGTTATTAATAATAA	1295
Query Match	54.7%; Score 726; DB 6; Length 30625;	24211	GTATTTTATTAATAATGTTGCCCATTTTATGAGTTATGATCAATTTGTTATTAATAATAA	24270
Best Local Similarity	99.4%; Pred. No. 0;	1296	GTTACAGATGTCA	1308
Matches	1146; Conservative 0; Mismatches 6; Indels 1; Gaps 1;	24271	GTTACAGATGTCA	24293
QY	157 TCCACATATTATGTCACACCTTCCCTGCCACAGGCGCGAGGACAGGGTGAG	23131	TCCACATATTATGTCACACCTTCCCTGCCACAGGCGCGAGGACAGGGTGAG	23190
Db	23131 TCCACATATTATGTCACACCTTCCCTGCCACAGGCGCGAGGACAGGGTGAG	217	GGTATACCCAAAGCTGTGACAGCCCAATTAGCCTTAAAGCAACTGACAGCAAGCCCTCC	276
QY	217 GGTATACCCAAAGCTGTGACAGCCCAATTAGCCTTAAAGCAACTGACAGCAAGCCCTCC	23191	GGTATACCCAAAGCTGTGACAGCCCAATTAGCCTTAAAGCAACTGACAGCAAGCCCTCC	23250
Db	23191 GGTATACCCAAAGCTGTGACAGCCCAATTAGCCTTAAAGCAACTGACAGCAAGCCCTCC	277	CTGGATGATCGAGTCCCGAGTCTGAAACAGAGTCCAGCAACCTCTTTCAGCCAG	336
QY	277 CTGGATGATCGAGTCCCGAGTCTGAAACAGAGTCCAGCAACCTCTTTCAGCCAG	23251	CTGGATGATCGAGTCCCGAGTCTGAAACAGAGTCCAGCAACCTCTTTCAGCCAG	23310
Db	23251 CTGGATGATCGAGTCCCGAGTCTGAAACAGAGTCCAGCAACCTCTTTCAGCCAG	337	GCCTCTGTGACCTGCTAGGTCAGAGGCTTCCAGAACAGTCTGTTGTAATTAGGACCC	396
QY	337 GCCTCTGTGACCTGCTAGGTCAGAGGCTTCCAGAACAGTCTGTTGTAATTAGGACCC	23311	GCCTCTGTGACCTGCTAGGTCAGAGGCTTCCAGAACAGTCTGTTGTAATTAGGACCC	23370
Db	23311 GCCTCTGTGACCTGCTAGGTCAGAGGCTTCCAGAACAGTCTGTTGTAATTAGGACCC	397	AAGCACT- GGGAGGGGCTGTTGGCTAGACCCCTTGTGAGACTTGGCACTATCTCAGTTA	455
QY	397 AAGCACT- GGGAGGGGCTGTTGGCTAGACCCCTTGTGAGACTTGGCACTATCTCAGTTA	23371	AAGCACTGGGAGGGGCTGTTGGCTAAACCCCTTGTGAGACTTGGCACTATCTCAGTTA	23430
Db	23371 AAGCACTGGGAGGGGCTGTTGGCTAAACCCCTTGTGAGACTTGGCACTATCTCAGTTA	456	GGATCTCTGCAGAAAACAAGACCACTTGTAGCTGGTTTAAATTAGACAAGATTACT	515
QY	456 GGATCTCTGCAGAAAACAAGACCACTTGTAGCTGGTTTAAATTAGACAAGATTACT	23431	GGATCTCTGCAGAAAACAAGACCACTTGTAGCTGGTTTAAATTAGACAAGATTACT	23490
Db	23431 GGATCTCTGCAGAAAACAAGACCACTTGTAGCTGGTTTAAATTAGACAAGATTACT	516	ACCTGGCCCTGCTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAACACACTCTGCTGAA	575
QY	516 ACCTGGCCCTGCTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAACACACTCTGCTGAA	23491	ACCTGGCCCTGCTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAACACACTCTGCTGAA	23550
Db	23491 ACCTGGCCCTGCTGGCTTGCAAAATTGTTGGAAGAGCTGGAGAACACACTCTGCTGAA	576	TTTCCAGGAACCTCCACGCGCAGATTCACTATGCTGTTGTGACCAAGAAAGCTGCCCC	635
QY	576 TTTCCAGGAACCTCCACGCGCAGATTCACTATGCTGTTGTGACCAAGAAAGCTGCCCC	23551	TTTCCAGGAACCTCCACGCGCAGATTCACTATGCTGTTGTGACCAAGAAAGCTGCCCC	23610
Db	23551 TTTCCAGGAACCTCCACGCGCAGATTCACTATGCTGTTGTGACCAAGAAAGCTGCCCC	636	ATCTGAGGAGCCATGTCGAGAAAGCTGCTGACTGACGACGAACTAGGCTCCCTCTGCCA	695
QY	636 ATCTGAGGAGCCATGTCGAGAAAGCTGCTGACTGACGACGAACTAGGCTCCCTCTGCCA	23611	ATCTGAGGAGCCATGTCGAGAAAGCTGCTGACTGACGACGAACTAGGCTCCCTCTGCCA	23670
Db	23611 ATCTGAGGAGCCATGTCGAGAAAGCTGCTGACTGACGACGAACTAGGCTCCCTCTGCCA	696	CGGTCCGCTGCCAGCCAATAGATGTCTCTGAGGCTGCCCCCTCTCCCACTTCACTCAGTTCC	755
QY	696 CGGTCCGCTGCCAGCCAATAGATGTCTCTGAGGCTGCCCCCTCTCCCACTTCACTCAGTTCC	23671	CGGTCCGCTGCCAGCCAATAGATGTCTCTGAGGCTGCCCCCTCTCCCACTTCACTCAGTTCC	23730
Db	23671 CGGTCCGCTGCCAGCCAATAGATGTCTCTGAGGCTGCCCCCTCTCCCACTTCACTCAGTTCC	756	CAAAATCTAAATTTTAAAGAGATCTGTTTGGGGGAACTTAAGTACAGTCCAGAACCTT	815
QY	756 CAAATCTAAATTTTAAAGAGATCTGTTTGGGGGAACTTAAGTACAGTCCAGAACCTT	23731	CAAAATCTAAATTTTAAAGAGATCTGTTTGGGGGAACTTAAGTACAGTCCAGAACCTT	23790
Db	23731 CAAATCTAAATTTTAAAGAGATCTGTTTGGGGGAACTTAAGTACAGTCCAGAACCTT	816	GGCTGCAAGGAGTCTGGGAAATGTCATTTCCCTAGAAGAAAGTTAGGGTGGGTGAGCA	875
QY	816 GGCTGCAAGGAGTCTGGGAAATGTCATTTCCCTAGAAGAAAGTTAGGGTGGGTGAGCA	23791	GGCTGCAAGGAGTCTGGGAAATGTCATTTCCCTAGAAGAAAGTTAGGGTGGGTGAGCA	23850
Db	23791 GGCTGCAAGGAGTCTGGGAAATGTCATTTCCCTAGAAGAAAGTTAGGGTGGGTGAGCA	876	AGCCCCACCTCGCTTTTCTGCCACAGCATCCCAATCGTGAAGAACTCCGGAGAGGGTGGA	935
QY	876 AGCCCCACCTCGCTTTTCTGCCACAGCATCCCAATCGTGAAGAACTCCGGAGAGGGTGGA			

Qy	921	TCGGAGAGGGTGGAGTCCACATCTAGGGTTGTCTGCCCTTGGCTCTATCCCTGGCCA	980
Db	1825	TCGGAGAGGGTGGAGTCCACATCTAGGGTTGTCTGCCCTTGGCTCTATCCCTGGCCA	1884
Qy	991	GAGGTGGAACTGGAGGAGTGGGTGCNAGACTGAGCTTAAATGTCCTCCCGGGCTTGAC	1040
Db	1885	GAGGTGGAACTGGAGGAGTGGGTGCNAGACTGAGCTTAAATGTCCTCCCGGGCTTGAC	1944
Qy	1041	TTTTCTTTTCTAGTCTCTGGGGCTAGATTCTCACTTGGGGTCTCTGACACAAACACACAT	1100
Db	1945	TTTTCTTTTCTAGTCTCTGGGGCTAGATTCTCACTTGGGGTCTCTGACACAAACACACAT	2004
Qy	1101	CCCAAAGTAGCCGAAGAGCTAAACACAGGGGGTCTTAAATAGGCTGCCCGGCCACCC	1160
Db	2005	CCCAAAGTAGCCGAAGAGCTAAACACAGGGGGTCTTAAATAGGCTGCCCGGCCACCC	2064
Qy	1161	GGGGCTCCCTTGGCCAAAGAAATGTGACCCCTACCCCAACCCCTCAACTACGAAATC	1220
Db	2065	GGGGCTCCCTTGGCCAAAGAAATGTGACCCCTACCCCAACCCCTCAACTACGAAATC	2124
Qy	1221	TGGGCCACCCGACGAGTATTTTATTTTAAATGTGGCCATTTTATCAGTATTATGTCAT	1280
Db	2125	TGGGCCACCCGACGAGTATTTTATTTTAAATGTGGCCATTTTATCAGTATTATGTCAT	1280
Qy	1281	TTGTATTAAATTTAAAGTTACAGA	1303
Db	2185	TTGTATTAAATTTAAAGTTACAGA	2207

RESULT	17
LOCUS	AX776017
DEFINITION	Sequence 287 from Patent WO03048202.
ACCESSION	AX776017
VERSION	AX776017.1 GI:32693735
KEYWORDS	.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 Matsuda, A. and Muramatsu, S.
AUTHORS	NF-kB activating gene
TITLE	Patent: WO 03048202-A 287 12-JUN-2003;
JOURNAL	Asahi Kasei Kabushiki Kaisha (JP)

```

location/Qualifiers
1..2207
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
571..1755
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE11708.1"
/db_xref="GI:32693736"
/translation="VACSLKDELLGICLSIYQDPVSLGCEHYFCRCRITETHWVQEQA
QGARDPCERFTFAPALAPSILKIANIYERYSFPDALLNARRAFCQAHDKVKLF
CLTDRALCLCFDEFAHGEHQVTDIDAFDELQDKDQALQDSREHTEALQLI
KQLAETKSSFTTITGEAFERLHRLRERKAMLEADTARTLITDEQKVQRY
SQLRKVQEGAIIQLERLAEDRHTFLAGVLSIRLKGKHTNLTVEDPSPKTYG
PIQYTLHWKLFQDIHPVPAALTDPTGAHORILSLDDCTIIVAYGNLHPQIQDPPKRF
DVEVSVLGSEATSSGGVHYEVVVVAEKTVQVIGLAHEAARSKGSIQIQNLGCKGVWEMS
F3"

```

	Query Match	34.1%	Score 452;	DB 6;	Length 2207;
	Best Local Similarity	99.8%;	Pred. No. 1.6e-263;		
	Matches 502;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	801	CAGATCCAGAACTTGGCTGTCGAAGGGAGCTCGGGAATATGTCATTTCCTTGAAGAGGAAGTT	860		
Db	1705	CAGATCCAGAACTTGGCTGTCGAAGGGAGCTCGGGAATATGTCATTTCCTTGAAGAGGAAGTT	1764		

QY	861	AGGTTGGGTGAGCAAGCCCACTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAAC	920
Db	1765	AGGTTGGGTGAGCAAGCCCACTGCGTTTTTCTGCCACAGCATCCAATCGTGAAGAAC	1824
QY	921	TCGGAGAGGGTGGAGTCCACATCTAGGGTTGTCTGCTCCCTCTGGCTCTATCCCTGCCCCA	980
Db	1825	TCGGAGAGGGTGGAGTCCACATCTAGGGTTGTCTGCTCCCTCTGGCTCTATCCCTGCCCCA	1884
QY	981	GAGGTGGAACTGCAGAGAGTGGGCTGCAAGACTGCAGCGCTAAATGTCTCCCGGCGCTTGAC	1040
Db	1885	GAGGTGGAACTGCAGAGAGTGGGCTGCAAGACTGCAGCGCTAAATGTCTCCCGGCGCTTGAC	1944
QY	1041	TTTTCTTTCTAGTCCTCGGGCCTAGATTTCTGCACTTTGGGCTCTCTGACACAAACACACCAT	1100
Db	1945	TTTTCTTTCTAGTCCTCGGGCCTAGATTTCTGCACTTTGGGCTCTCTGACACAAACACACCAT	2004
QY	1101	CCCAAAGTAGCCGAAGAGCTAAACACAGGGGGTCTTAAATGGCTGCCCGCCGACACC	1160
Db	2005	CCCAAAGTAGCCGAAGAGCTAAACACAGGGGGTCTTAAATGGCTGCCCGCCGACACC	2064
QY	1161	GGGCCTCCCTTGGGCAAAAGAAATGTGACGCCCTACCCCAACCCCTTCAACTACCGAAATC	1220
Db	2065	GGGCCTCCCTTGGGCAAAAGAAATGTGACGCCCTACCCCAACCCCTTCAACTACCGAAATC	2124
QY	1221	TGGGCCACCCACAGCATTTTTTATTTTAAATGGCTGGCCCATTTTATGAGTTATGATCAAT	1280
Db	2125	TGGGCCACCCACAGCATTTTTTATTTTAAATGGCTGGCCCATTTTATGAGTTATGATCAAT	2184
QY	1281	TTGTATTAATTTAAAGTTACAGA	1303
Db	2185	TTGTATTAATTTAAAGTTACAGA	2207

RESULT	18
AX775829	
LOCUS	AX775829
DEFINITION	Sequence 99 from Patent WO03048202.
ACCESSION	AX775829
VERSION	AX775829..1 GI:32693547
KEYWORDS	.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 Matsuda A. and Muramatsu S. AUTHORS TITLE NF-kB activating gene Patent: WO 03048202-A 99 12-JUN-2003; Asahi Kasei Kabushiki Kaisha (Jp) JOURNAL

```

FEATURES             source
Location/Qualifiers
1..2246
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
138..485
  /note="unnamed protein product"
/codon_start=1
/protein_id="CAE11614.1"
/db_xref="GI:32693548"
/translation="MYGSAWLPPPPGFFIPFSPSPHRPHGISTPLRAPPSGIPFS
QLLSRTAPPRIPSPGSPSLRSQPPPSAHRLELPTSLRPSAGVPSLSRAPLPR
VTLPLDPTPSLLR"

```

Query Match	34.1%	Score 452;	DB 5;	Length 2246;
Best Local Similarity	99.6%	Pred. No. 1.6e-263;		
Matches 502;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	801	CAGATCCGAGACCTTGGCTGCAAGGGAGCTCGGGAATGTCATTTCCTTCCCTAGAGGAAGTT	860	
Db	1744	CAGATCCGAGACCTTGGCTGCAAGGGAGCTCGGGAATGTCATTTCCTTCCCTAGAGGAAGTT	1803	

QY 1065 GATTCTGCACTTGGGGTCTCTGACACAAACACCAATCCCAAGTAGCCGGAAGACTAAA 1124
Db 243 GATTCTGCACTTGGGGTCTCTGACACAAACACCAATCCCAAGTAGCCGGAAGACTAAA 184
QY 1125 CACAGGGGTTCTTAATGCTGCTCCCGCCCGCCCGGCTCCCTTTGGGCAAAAGGAAT 1184
Db 183 CACAGGGGTTCTTAATGCTGCTCCCGCCCGCCCGGCTCCCTTTGGGCAAAAGGAAT 124
QY 1185 TGTGAGCCCTACCCCAACCTTCAATACACAGAACTCTGGGCGCCACCCAGCAGTATTTT 1244
Db 123 TGTGAGCCCTACCCCAACCTTCAATACACAGAACTCTGGGCGCCACCCAGCAGTATTTT 64
QY 1245 TTTAAATGTTGCCATTTTATGAGTTATGATCAATTTGTTATTAATTAAGTTACAGAT 1304
Db 63 TTTAAATGTTGCCATTTTATGAGTTATGATCAATTTGTTATTAATTAAGTTACAGAT 4
QY 1305 GTC 1307
Db 3 GTC 1
RESULT 20
BD153810/c
LOCUS BD153810 557 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD153810
VERSION BD153810.1 GI:27859568
KEYWORDS JP 2002191363-A/8653.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 557)
AUTHORS Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE Primers for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002191363-A 8653 09-JUL-2002;
COMMENT HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/8653
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10,
PC C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT source 1..557
FT /organism='Homo sapiens (human)'.
FEATURES
source
1..557
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN
Query Match 31.9%; Score 423; DB 6; Length 557;
Best Local Similarity 100.0%; Pred. No. 7.2e-246;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 885 TGGGTTTCTTCCACAGCATCAATCGTGAAGAACTCGGAGAGGTTGGAGTCCACATC 944
Db 423 TGGGTTTCTTCCACAGCATCAATCGTGAAGAACTCGGAGAGGTTGGAGTCCACATC 364
QY 945 TAGGGTTGTCTGCCCTTGGCTCTATCCCTGCCAGAGTGGAACTGGAGGAGTGGGC 1004
Db 363 TAGGGTTGTCTGCCCTTGGCTCTATCCCTGCCAGAGTGGAACTGGAGGAGTGGGC 304

QY 861 AGGTGGTGGAGCAAGCCCACTCGTTTCTGCGCAGAGCATCAATCGTGAAGAAC 920
Db 1804 AGGTGGTGGAGCAAGCCCACTCGTTTCTGCGCAGAGCATCAATCGTGAAGAAC 1863
QY 921 TCGGAGAGGGTGGAGTCCACATCTAGGTTGTCTGCTCCCTTGGCTCTATCCCTGCCCA 980
Db 1864 TCGGAGAGGGTGGAGTCCACATCTAGGTTGTCTGCTCCCTTGGCTCTATCCCTGCCCA 1923
QY 981 GAGGTGGAACTGGAGAGTGGGTGCAAGACTGAGCTAAATGTCTCCCGGCTTGAC 1040
Db 1924 GAGGTGGAACTGGAGAGTGGGTGCAAGACTGAGCTAAATGTCTCCCGGCTTGAC 1983
QY 1041 TTTTCTTCTAGTCTGGGGCTAGATCTGCACTTGGGTCCTGACACACACACCAT 1100
Db 1984 TTTTCTTCTAGTCTGGGGCTAGATCTGCACTTGGGTCCTGACACACACACCAT 2043
QY 1101 CCCAAAGTAGCGGGAAGCTAAACACAGAGGGGTTCTTAAATGCTGCCCGGCCACCC 1160
Db 2044 CCCAAAGTAGCGGGAAGCTAAACACAGAGGGGTTCTTAAATGCTGCCCGGCCACCC 2103
QY 1161 GGGCTCCCTTGGGCAAGAAATGTGAGCCCTACCCCAACCTTCACTACCAGATC 1220
Db 2104 GGGCTCCCTTGGGCAAGAAATGTGAGCCCTACCCCAACCTTCACTACCAGATC 2163
QY 1221 TGGGCCACCCAGCAGTATTTTATTAATGTTGCCATTTTATGAGTTATGATCAAT 1280
Db 2164 TGGGCCACCCAGCAGTATTTTATTAATGTTGCCATTTTATGAGTTATGATCAAT 2223
QY 1281 TTGTATTAAATTAAGTTACAGA 1303
Db 2224 TTGTATTAAATTAAGTTACAGA 2246
RESULT 19
AX873748/c
LOCUS AX873748 557 bp DNA linear PAT 17-DEC-2003
DEFINITION Sequence 8653 from Patent EP1074617.
ACCESSION AX873748
VERSION AX873748.1 GI:40028540
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE Primers for synthesizing full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 8653 07-FEB-2001;
COMMENT Research Association for Biotechnology (JP)
Location/Qualifiers
FT source 1..557
FT /organism='Homo sapiens'
/mol_type='unassigned DNA'
/db_xref='taxon:9606'
ORIGIN
Query Match 31.9%; Score 423; DB 6; Length 557;
Best Local Similarity 100.0%; Pred. No. 7.2e-246;
Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 885 TGGGTTTCTTCCACAGCATCAATCGTGAAGAACTCGGAGAGGTTGGAGTCCACATC 944
Db 423 TGGGTTTCTTCCACAGCATCAATCGTGAAGAACTCGGAGAGGTTGGAGTCCACATC 364
QY 945 TAGGGTTGTCTGCCCTTGGCTCTATCCCTGCCAGAGTGGAACTGGAGGAGTGGGC 1004
Db 363 TAGGGTTGTCTGCCCTTGGCTCTATCCCTGCCAGAGTGGAACTGGAGGAGTGGGC 304
QY 1005 TGAAGACTGAGCCTAAATGTCTCCCGGCTTGACATTTTCTTCTAGTCTGGGGCTTA 1064
Db 303 TGAAGACTGAGCCTAAATGTCTCCCGGCTTGACATTTTCTTCTAGTCTGGGGCTTA 244

Qy	1005	TGCAAGACTGAGCCTAAATGTCCTCCCGGCTTGCATTTCTTCTAGTCTCTGGGGCCTA	1064
Db	303	TGCAAGACTGAGCCTAAATGTCCTCCCGGCTTGCATTTCTTCTAGTCTCTGGGGCCTA	244
Qy	1065	GATTCTGCACCTTGGGGTCTCTGACACAACACACACCATCCCAAAGTAGCCGGAAGAGCTAAA	1124
Db	243	GATTCTGCACCTTGGGGTCTCTGACACAACACACCATCCCAAAGTAGCCGGAAGAGCTAAA	184
Qy	1125	CACAGGGGGTCTTTAAAATGGCTGCCCCCGGCCACCCGGGGCTCCCTTTGGGCAAAAGGAAT	1184
Db	183	CACAGGGGGTCTTTAAAATGGCTGCCCCCGGCCACCCGGGGCTCCCTTTGGGCAAAAGGAAT	124
Qy	1185	TGTCAGCGCCTACCCCAACCCCTTCAACTACACAGAATCTGGGGCCACCCCGACGAGTATTTT	1244
Db	123	TGTCAGCGCCTACCCCAACCCCTTCAACTACACAGAATCTGGGGCCACCCCGACGAGTATTTT	64
Qy	1245	TTTAAAATGTTGCCATTTTATGAGTTATGATCAATTTGTTATTAATTTAAAGTTACAGAT	1304
Db	63	TTTAAAATGTTGCCATTTTATGAGTTATGATCAATTTGTTATTAATTTAAAGTTACAGAT	4
Qy	1305	GTC 1307	
Db			
	3	GTC 1	
RESULT 21			
AK122896			
LOCUS		3148 bp	mRNA
DEFINITION	AK122896		linear
		Homo sapiens cDNA FLJ16558 fis, clone SYN04001153, weakly similar	
		to Zinc-binding protein A33.	
ACCESSION	AK122896		
VERSION	AK122896.1	GI:34528340	
KEYWORDS		oligo capping; fis (full insert sequence).	

[illegible]

RESULT 22	GI5084	human STS SHGC-15366, sequence tagged site.	250 bp	DNA	linear	STS
LOCUS	GI5084					
DEFINITION	GI5084					
ACCESSION	GI5084.1	GI:1131847				
VERSION	STS; STS sequence; primer; sequence tagged site.					
KEYWORDS	Homo sapiens (human)					
SOURCE	Homo sapiens					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eut					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Ho					
JOURNAL	1 (bases 1 to 250)					
COMMENT	Myers, R.M.					
	Unpublished (1995)					
	Contact: Richard M. Myers					
	Stanford Human Genome Center (SHGC)					
	Stanford University School of Medicine					
	Department of Genetics, M-344, Stanford, CA 94305, USA					
	Tel: 4157259687					
	Fax: 4157259689					
	Email: myers@shgc.stanford.edu					
	Primer A: AGTTCAGGGGTGGGTAGG					
	Primer B: CACTGGGGTCTCTGACACA					
	STS size: 140					
	PCR Profile:					
	Initial incubation: 94 degrees C for 90 seconds					
	Denaturation: 94 degrees C for 15 seconds					
	Annealing: 62 degrees C for 23 seconds					
	Polymerization: 72 degrees C for 30 seconds					
	PCR Cycles: 30					
	Thermal Cycler: Perkin Elmer 9600					
	Protocol:					
	Template: 25 ng					
	Primer: each 1 uM					
	dNTPs: each 200 uM					
	Tag Polymerase: 0.05 units/ul					
	Total Vol: 10 ul					
	Buffer:					
	MgCl2: 2.5 mM					
	KCl: 50 mM					

```

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
  source
    1. .250
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /map="1"
    109. .248
    109. .128
    complement(229. .248)
ORIGIN
  Query Match      4.7%; Score 62; DB 11; Length 250;
  Best Local Similarity 100.0%; Pred. No. 1.4e-25;
  Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1213 CCAGAACTGGCCACCCACGAGTATTTTATTAATGTTGCCATTTTATGAGTTA 1272
      |||||||
Db 107 CCAGAACTGGCCACCCACGAGTATTTTATTAATGTTGCCATTTTATGAGTTA 48

Qy 1273 TG 1274
      ||
Db 47 TG 46

RESULT 23
AX388847
LOCUS AX388847 614 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 3775 from Patent WO0214500.
ACCESSION AX388847
VERSION AX388847.1 GI:19581977
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Escobedo, J., Garcia, P.D., Sudduth-Klinger, J., Reinhard, C.,
Randoz, F., Lamson, G., Scott, E.M., Zhang, G., Kassam, A., Pot, D. and
Labat I.
TITLE Human genes and gene expression products
JOURNAL Patent: WO 0214500-A 3775 21-FEB-2002;
CHIRON CORPORATION (US) ; Hyseq Inc. (US)
FEATURES
  source
    1. .614
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"
ORIGIN
  Query Match      2.2%; Score 29; DB 6; Length 614;
  Best Local Similarity 100.0%; Pred. No. 2.1e-05;
  Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1299 ACAGATGTCAAAAAAAAAAAAAAAAAAAAA 1327
      |||||||
Db 585 ACAGATGTCAAAAAAAAAAAAAAAAAAAAA 613

RESULT 24
AX771586/c
LOCUS AX771586 1652 bp DNA linear PAT 02-JUL-2003
DEFINITION Sequence 299 from Patent WO03004646.
ACCESSION AX771586
VERSION AX771586.1 GI:32438384
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

1
O'Mahony, D.J., Byrne, D., Brayden, D., Lambkin, I. and Higgins, I.
Genetic analysis of Peyer's patches and M cells and methods and
compositions targeting Peyer's patches and M cell receptors
Patent: WO 03004646-A 299 16-JAN-2003;
ELAN CORPORATION, PLC (IE)
FEATURES
  Location/Qualifiers
    source
      1. .1652
      /organism="Homo sapiens"
      /mol_type="unassigned DNA"
      /db_xref="taxon:9606"
ORIGIN
  Query Match      2.2%; Score 29; DB 6; Length 1652;
  Best Local Similarity 100.0%; Pred. No. 2.3e-05;
  Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1299 ACAGATGTCAAAAAAAAAAAAAAAAAAAAA 1327
      |||||||
Db 1474 ACAGATGTCAAAAAAAAAAAAAAAAAAAAA 1446

RESULT 25
AC019255/c
LOCUS AC019255 162959 bp DNA linear HTG 24-AUG-2002
DEFINITION Homo sapiens chromosome 12 clone RP11-133N3 map 12, WORKING DRAFT
SEQUENCE, 17 unordered pieces.
ACCESSION AC019255
VERSION AC019255.3 GI:10045315
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 12, clone RP11-133N3
JOURNAL Unpublished
AUTHORS
  2 (bases 1 to 162959)
  Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
  Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F.,
  Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G., Castle, A.,
  Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
  DeArellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,
  Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,
  Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
  Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
  Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K.,
  Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
  McPheeters, R., Meldrim, J., Meneus, L., Morrow, J., Naylor, J.,
  Norman, C.H., O'Connor, T., O'Donnell, P., Olivari, T.M., Peterson, K.,
  Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
  Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
  Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
  Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,
  Zimmer, A. and Zody, M.
  Direct Submission
  Submitted (31-DEC-1999) Whitehead Institute/MIT Center for Genome
  Research, 320 Charles Street, Cambridge, MA 02141, USA
  3 (bases 1 to 162959)
  Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
  Anderson, S., Barna, N., Bastien, V., Beda, F., Boguslavskiy, L.,
  Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,
  Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
  DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P.,
  FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,
  Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L.,
  Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., LaRocque, K.,
  Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G.,
  Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McKernan, K.,
  McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
  Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,
  O'Donnell, P., O'Neill, D., Olivari, T.M., Oliver, J., Peterson, K.,

```

Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Sogniez,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Sep 9, 2000 this sequence version replaced gi:16759192.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www.seq.wi.mit.edu

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L2578

Center clone name: 133 N_3

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 154857 bases at least Q40

Consensus quality: 158391 bases at least Q30

Consensus quality: 159656 bases at least Q20

Insert size: 172000; agarose-fp

Insert size: 161359; sum-of-contigs

Quality coverage: 5.1 in Q20 bases; agarose-fp

Quality coverage: 5.4 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 30551: contig of 30551 bp in length
* 30552 30651: gap of 100 bp
* 30652 33510: contig of 2859 bp in length
* 33510 33610: gap of 100 bp
* 33610 33611: contig of 2871 bp in length
* 33611 36481: gap of 100 bp
* 36481 36482 39763: contig of 3182 bp in length
* 39763 39764 39863: gap of 100 bp
* 39863 45400: contig of 5537 bp in length
* 45400 45500: gap of 100 bp
* 45500 49120: contig of 3620 bp in length
* 49120 49221: gap of 100 bp
* 49221 53633: contig of 4413 bp in length
* 53633 53734: gap of 100 bp
* 53734 58228: contig of 4495 bp in length
* 58228 58229 58328: gap of 100 bp
* 58328 63731: contig of 5403 bp in length
* 63731 63831: gap of 100 bp
* 63831 69278: contig of 5447 bp in length
* 69278 69379: gap of 100 bp
* 69379 76967: contig of 7589 bp in length
* 76967 77068 77067: gap of 100 bp
* 77067 86466: contig of 9399 bp in length
* 86466 86566: gap of 100 bp
* 86566 97145: contig of 10579 bp in length
* 97145 97245: gap of 100 bp
* 97245 110405: contig of 13160 bp in length
* 110405 110506: gap of 100 bp
* 110506 135582: contig of 25077 bp in length
* 135582 161840: gap of 100 bp
* 161840 161840: contig of 26158 bp in length

```

* 161841 161940: gap of 100 bp
* 161941 162959: contig of 1019 bp in length.

FEATURES

source

Location/Qualifiers
1. .162959
/organism="Homo sapiens"
/mol type="genomic DNA"
/db xref="taxon:9606"
/chromosome="12"
/map="12"

/clone="RP11-133N3"
/clone lib="RPC1-11 Human Male BAC"

1. .30551

/note="assembly_fragment"

clone_end:SP6

vector_side:left"

30652..33510

/note="assembly_fragment"

33611..36481

/note="assembly_fragment"

36582..39763

/note="assembly_fragment"

39864..45400

/note="assembly_fragment"

45501..49120

/note="assembly_fragment"

49221..53633

/note="assembly_fragment"

53734..58228

/note="assembly_fragment"

58329..63731

/note="assembly_fragment"

63832..69278

/note="assembly_fragment"

69379..76967

/note="assembly_fragment"

77068..86466

/note="assembly_fragment"

86567..97145

/note="assembly_fragment"

97246..110405

/note="assembly_fragment"

110506..135582

/note="assembly_fragment"

135683..161840

/note="assembly_fragment"

161941..162959

/note="assembly_fragment"

clone_end:T7

vector_side:right"

ORIGIN

Query Match 2.2%; Score 29; DB 2; Length 162959;

Best Local Similarity 100.0%; Pred. No. 3.3e-05;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 ACAGATGTCACAAAAA 1327

Db 101907 ACAGATGTCACAAAAA 101879

=====

=====

RESULT 26

AC022080

LOCUS

DEFINITION

Homo sapiens 12 BAC RP11-820K3 (Roswell Park Cancer Institute Human BAC Library) complete sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

1

(bases 1 to 178217)

AC022080

Homo sapiens

178217 bp

DNA

linear

PRI 15-MAR-2003

<p>AUTHORS</p> <p>Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amarantunga, H.C., Are, J.R., Ayele, M., Banks, T., Barbria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Liao, J., Liu, J., Liu, W., Louie, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Marandel, I., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Merscher, S., Metzker, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojebokan, I., Rolfe, M., Ruiz, S., Savary, G., Scher, S., Scott, G., Shen, H., Shim, C., Shooshtari, N., Sisson, L., Sodergren, E., Sonaikar, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, R., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Kucherlapati, R., Weinstein, G. and Gibbs, R.</p> <p>Direct Submission</p> <p>Unpublished</p> <p>2 (bases 1 to 178217)</p> <p>Worley, K.C.</p> <p>Direct Submission</p> <p>Submitted (26-JAN-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA</p> <p>3 (bases 1 to 178217)</p> <p>Worley, K.C.</p> <p>Direct Submission</p> <p>Submitted (24-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA</p> <p>4 (bases 1 to 178217)</p> <p>Worley, K.C.</p> <p>Direct Submission</p> <p>Submitted (29-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA</p> <p>5 (bases 1 to 178217)</p> <p>Worley, K.C.</p> <p>Direct Submission</p> <p>Submitted (30-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA</p> <p>6 (bases 1 to 178217)</p> <p>Worley, K.C.</p> <p>Direct Submission</p> <p>Submitted (30-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One</p>	<p>REFERENCE</p> <p>AUTHORS</p> <p>Worley, K.C.</p> <p>Direct Submission</p> <p>Submitted (15-MAR-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA</p> <p>On Jul 29, 2002 this sequence version replaced gi:21914534.</p> <p>INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu</p> <p>CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.</p> <p>ANNOTATION OF FEATURES:</p> <p>STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts</p> <p>Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.</p> <p>Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.</p> <p>SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.</p> <p>QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht</p> <p>misc_feature</p> <p>repeat_region</p> <p>repeat_region</p> <p>repeat_region</p> <p>misc_feature</p> <p>repeat_region</p> <p>misc_feature</p> <p>STS</p> <p>repeat_region</p>	<p>Location/Qualifiers</p> <p>1..178217</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="genomic DNA"</p> <p>/db_xref="taxon:9606"</p> <p>/chromosome="12"</p> <p>/clone="RP11-820K3"</p> <p>/complement(1..2015)</p> <p>/note="overlaps bases 1..2004 of clone AC068792"</p> <p>/function="clone overlap"</p> <p>/complement(119..424)</p> <p>/rpt_family="AluY"</p> <p>564..665</p> <p>/rpt_family="U13"</p> <p>/complement(760..1052)</p> <p>/rpt_family="AluDb"</p> <p>1583</p> <p>/note="polymorphism G/T AC022080/AC068792"</p> <p>/function="polymorphic site"</p> <p>/complement(1713..2020)</p> <p>/rpt_family="AluSg"</p> <p>1713</p> <p>/note="polymorphism T/A AC022080/AC068792"</p> <p>/function="polymorphic site"</p> <p>2073..2196</p> <p>/standard_name="RH92710"</p> <p>2261..2296</p> <p>/rpt_family="AT-rich"</p>
---	---	---

```

repeat_region      2943..3072
                    /rpt_family="FLAM_C"
repeat_region      3743..3775
                    /rpt_family="AT_rich"
repeat_region      4258..4559
                    /rpt_family="AluSx"
repeat_region      complement(4640..4683)
                    /rpt_family="MIR"
repeat_region      complement(4840..5140)
                    /rpt_family="AluSg1"
repeat_region      complement(5211..5517)
                    /rpt_family="AluY"
repeat_region      6138..6443
                    /rpt_family="AluY"
repeat_region      6446..6747
                    /rpt_family="AluSg"
repeat_region      7202..7489
                    /rpt_family="AluSx"
repeat_region      complement(8107..8411)
                    /rpt_family="AluY"
repeat_region      9329..9628
                    /rpt_family="AluSp"
repeat_region      complement(9687..9990)
                    /rpt_family="AluSc"
repeat_region      complement(11137..11423)
                    /rpt_family="AluSg"
repeat_region      complement(11476..11566)
                    /rpt_family="AluSg/X"
repeat_region      complement(11882..12185)
                    /rpt_family="MER46C"
repeat_region      13075..13377
                    /rpt_family="AluSg"
repeat_region      13378..13686
                    /rpt_family="AluSx"
repeat_region      13808..13903
                    /rpt_family="L2"
repeat_region      complement(14066..14365)
                    /rpt_family="AluSg"
repeat_region      14652..14782
                    /rpt_family="L2"
repeat_region      complement(14951..15260)
                    /rpt_family="AluSx"
repeat_region      complement(15772..16062)

Query Match      2.2%; Score 29; DB 9; Length 178217;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 ACAGATGTCACAAAAA...AAAAA 1327
DB 165508 ACAGATGTCACAAAAA...AAAAA 165536

RESULT 27
AC090496/c
LOCUS      Mus musculus clone rp23-469n6 map 14 strain C57BL/6J, complete
DEFINITION      Mus musculus clone rp23-469n6 map 14 strain C57BL/6J, complete
ACCESSION      AC090496
VERSION      HTG.
KEYWORDS      Mus musculus (house mouse)
SOURCE      Mus musculus
ORGANISM      Mus musculus
REFERENCE      1 (bases 1 to 180338)
AUTHORS      Wu, H., Hu, P., Yang, L., Prescott, A. and Roe, B.A.
TITLE      Mus musculus BAC Clone rp23-469n6
JOURNAL      Unpublished
REFERENCE      2 (bases 1 to 180338)
AUTHORS      Wu, H., Hu, P., Yang, L., Prescott, A. and Roe, B.A.
TITLE      Direct Submission
JOURNAL      Submitted (24-FEB-2001) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA

The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 180338)
AUTHORS      Wu, H., Hu, P., Yang, L., Prescott, A. and Roe, B.A.
TITLE      Direct Submission
JOURNAL      Submitted (18-JUN-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
4 (bases 1 to 180338)
AUTHORS      Wu, H., Hu, P., Yang, L., Prescott, A. and Roe, B.A.
TITLE      Direct Submission
JOURNAL      Submitted (20-JUN-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On Jun 20, 2003 this sequence version replaced gi:31880147.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
FEATURES             Location/Qualifiers
     source            1..180338
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /map="14"
                     /clone="rp23-469n6"
                     /clone_lib="RPCI - 23 Female (C57BL/6J) Mouse BAC Library"
ORIGIN
Query Match      2.2%; Score 29; DB 10; Length 180338;
Best Local Similarity 100.0%; Pred. No. 3.3e-05;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 812 CCTTGCTGCAAGGAGTCTGGGAAATGT 840
DB 73352 CCTTGCTGCAAGGAGTCTGGGAAATGT 73324

RESULT 28
BC060953
LOCUS      Mus musculus cDNA clone MGC:74041 IMAGE:6774045, complete cds.
DEFINITION      Mus musculus cDNA clone MGC:74041 IMAGE:6774045, complete cds.
ACCESSION      BC060953
VERSION      BC060953.1 GI:38511618
KEYWORDS      MGC.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
REFERENCE      1 (bases 1 to 723)
AUTHORS      Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Bhat, N.K.,
Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullen, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, J.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, J., Whitting, M., Madan, A., Young, A.C., Shchepochko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Small, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL

```



```

MEDLINE
PUBMED
REFERENCE
2. (bases 1 to 723)
AUTHORS
Strausberg, R.
TITLE
Direct Submission
JOURNAL
Submitted (03-NOV-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT
Contact: MGC help desk
Email: gcapps-remail.nih.gov
Tissue Procurement: Dr. Jonathan Kuo, NIMH
cDNA Library Preparation: Michael Brownstein / Ted Usdin
Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.sbgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 53 Row: h Column: 10
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 27734083.
FEATURES
source
1. .723
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="MGC:74041 IMAGE:6774045"
/tissue_type="Testis, mouse"
/clone_lib="NTH MGC_169"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
21..563
/codon_start=1
/product="Unknown (protein for MGC:74041)"
/protein_id="AAH60953.1"
/db_xref="GI:38511619"
/translation="MARVAKVHVSRAATAVAATAKAKRSKLKTKAKRSKLKTKAAK
RFLKLNKPNRSKLPKSHSLIHSFTRSPSCGCHCCCHCCLHSPSPYKSPFKLT
KKGQSLRRRIKRAIEIRLMQSHLRSQKLIELEPSEIIVAFSHENNVNVEP
QEVPPCLDSDFPNVDUASF"
CDS
2.1% Score 28; DB 10; Length 723;
Query Match
Best Local Similarity 100.0%; Pred. No. 8 6e-05;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
ORIGIN
QY 1300 CAGATGTCAAAAAAAAAAAAAAAAAAAAAA 1327
| | | | | | | | | | | | | | | | | | | |
Db 686 CAGATGTCAAAAAAAAAAAAAAAAAAAAAA 713
| | | | | | | | | | | | | | | | | | | |
RESULT 29
AC004104
LOCUS
DEFINITION
Homo sapiens 12q24 PAC RPCI5-942N13 (Roswell Park Cancer Institute
Human PAC library) complete sequence.
ACCESSION
AC004104
VERSION
AC004104.1 GI:3219327
KEYWORDS
HTG.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. (bases 1 to 136037)
AUTHORS
Muzny, D., Arenson, A.D., Adams, C., Bunac, C., Carvelli, K., Chang, J.,
Chacko, J., Chen, J., Ding, Y., Dugan, S., Durbin, J., Forcum, J.,
Ganesh, R., Garcia, C., Goodman, M., Gorrell, J.H., Haywood, M.,
Hernandez, J., Jackson, L., Jin, S., Kampal, R., Karpathy, S., Kovar, C.,
Lau, S., Leal, B., Lee, B., Li, Y., Lichtarge, O., Liu, W., Logan, O.,
Lu, J., Ly, T., Marondel, I., Martinez, C., Merscher, S., Montgomery, K.,
Oswal, G., Perez, L., Rashid, N.D., Renault, B., Rowland, K., Savage, L.,
Scherer, S.E., Shen, H., Simon, M., Stovall, K., Timms, K.M., Todd, J.,
Vo, Q., Williamson, A., Worley, K.C., Yu, W., Kucherlapati, R.,
Nelson, D. and Gibbs, R.A.
TITLE
Direct Submission
JOURNAL
Unpublished
REFERENCE
2. (bases 1 to 136037)
AUTHORS
Worley, K.C.
TITLE
Direct Submission
JOURNAL
Submitted (03-FEB-1998) Molecular and Human Genetics, Baylor
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE
3. (bases 1 to 136037)
AUTHORS
Worley, K.C.
TITLE
Direct Submission
JOURNAL
Submitted (13-JUN-1998) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jun 13, 1998 this sequence version replaced gi:3108008.
Sequencing is completed to a minimum standard of double strand
coverage with a minimum of 2 clones and 2 reads with no ambiguities
or 2 chemistries with a minimum of 2 clones and 3 reads with no
ambiguities. If the sequence quality does not meet this standard,
it will be indicated in the annotation.
The repeat regions shown were identified using RepeatMasker by
Adrian Smit.
Sequence similarities were identified using Powerblast by Jinghui
Zhang.
Exon/Intron boundaries of identified genes were chosen if there
were canonical splice junctions that maintained sequence continuity
across the splice junctions.
FEATURES
source
1. .136037
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/map="12q24"
/clone="RPCI5-942N13"
/clone_lib="Roswell Park Cancer Institute Human PAC
library"
1..113
/rpt_family="L1PA11"
complement(353..498)
/rpt_family="MIR"
499..805
/rpt_family="AluSp"
complement(806..826)
/rpt_family="AT-rich"
complement(905..995)
/rpt_family="(GAAA)n"
complement(1848..2156)
/rpt_family="AluX"
complement(2299..2519)
/rpt_family="MIR"
complement(2641..2806)
/rpt_family="MIR"
2813..2994
/rpt_family="MER53"
complement(3189..3307)
/rpt_family="MIR"
3312..3522
/rpt_family="MER20"
complement(3538..3656)
/rpt_family="MIR"
complement(4089..4222)
/rpt_family="LTR16C"

```

```

repeat_region 4234..4764
/rpt family="L1PA16"
complement(4786..4877)
/rpt family="LTR16C"
complement(5068..5103)
/rpt family="L2"
complement(5069..5121)
/rpt family="MIR"
5481..5514
/rpt family="AT_rich"
6376..6597
/rpt family="MIR"
7116..7179
/rpt family="MIR"
complement(7197..7285)
/rpt family="L2"
7709..7851
/rpt family="MIR"
complement(7876..8303)
/rpt family="L2"
8308..8386
/rpt family="MIR"
complement(8724..8818)
/rpt family="L1MB6"
8911..9237
/rpt family="AluJ0"
complement(9292..9586)
/rpt family="AluSx"
9912..9980
/rpt family="MIR"
10103..10156
/rpt family="L2"
complement(10161..10506)
/rpt family="MIR1B"
complement(10693..10846)
/rpt family="MIR"
11030..11189
/rpt family="L2"
complement(11207..11503)
/rpt family="AluJ0"
11557..11865
/rpt family="AluJb"
complement(12046..12642)
/rpt family="L1MB6"
12676..12962
/rpt family="AluJ0"
complement(13150..13192)
/rpt family="(TGAA)n"
complement(13193..13295)
/rpt family="L2"
13300..13582
/rpt family="MER7B"
13587..13866
/rpt family="AluSx"
13867..14110
/rpt family="MER7B"
complement(14121..14228)
/rpt family="L1PA14"
14233..14618
/rpt family="MER7B"
complement(14845..14891)
/rpt family="(GAAA)n"
complement(15165..15872)
/rpt family="TIGGER1"
complement(16050..16351)
/rpt family="AluSx"
complement(16688..16724)
/rpt family="L1PA15"
complement(16947..17820)
/rpt family="L1PA16"
17842..18169
/rpt family="MER7B"
complement(18193..18536)

```

```

repeat_region /rpt family="L2"
complement(18619..18653)
/rpt family="AT_rich"
complement(18654..18955)
/rpt family="AluSx"
complement(18956..19180)
/rpt family="L2"
complement(19502..20558)
/rpt family="L1MC3"
20563..20675
/rpt family="(TA)n"
complement(20693..20873)
/rpt family="L1MD3"
20994..21182
/rpt family="AluJ0"
complement(21190..21321)
/rpt family="MIR"
21514..21604
/rpt family="MIR"
21692..21751
/rpt family="MIR"
21749..21981
/rpt family="MIR"
complement(22820..22879)
/rpt family="(TGA)n"
complement(22891..22977)
/rpt family="MIR"
23027..23114
/rpt family="MER58B"
23169..23307
/rpt family="MER58A"
complement(23539..23573)
/rpt family="(CA)n"
23573..23612
/rpt family="(GA)n"
complement(23695..23996)
/rpt family="AluJ0"
complement(24725..24939)
/rpt family="L2"
25378..25713
/rpt family="LTR16C"

```

Query Match 2.1%; Score 28; DB 9; Length 136037;
 Best Local Similarity 100.0%; Pred. No. 0.00013;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 813 CTTGGCTGCAAGGGAGTCTGGGAAATGT 840
 |||||
 Db 65773 CTTGGCTGCAAGGGAGTCTGGGAAATGT 65800

RESULT 30
 AC024316
 LOCUS
 DEFINITION Homo sapiens clone RP11-2608, WORKING DRAFT SEQUENCE, 17 unordered
 pieces.
 AC024316
 VERSION HTG; HTGS PHASE1; HTGS_DRAFT.
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 137648)
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Balgwin,J., Barna,N., Beda,F., Boguslavskiy,L.,
 Bouckgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
 Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
 DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M.,
 Unpublished
 JOURNAL
 TITLE
 AUTHORS
 REFERENCE 2 (bases 1 to 137648)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
 Anderson,S., Balgwin,J., Barna,N., Beda,F., Boguslavskiy,L.,
 Bouckgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
 Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
 DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M.,

Fenestor, J., Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Landers, T., Largocque, K., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivari, T. M., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testave, S., Theodore, J., Tirrell, A., Travers, M., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (28-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 26, 2000 this sequence version replaced gi:10198462.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L4672
Center clone name: 26_Q 8
----- Summary Statistics
Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 128694 bases at least Q40
Consensus quality: 133079 bases at least Q30
Consensus quality: 135006 bases at least Q20
Insert size: 137000; agarose-fp
Insert size: 136048; sum-of-contigs
Quality coverage: 4.7 in Q20 bases; agarose-fp
Quality coverage: 4.8 in Q20 bases; sum-of-contigs

***** NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 3290: contig of 3290 bp in length
3291 3390: gap of 100 bp
3391 5203: contig of 1813 bp in length
5204 5303: gap of 100 bp
5304 7364: contig of 2061 bp in length
7365 7464: gap of 100 bp
7465 10939: contig of 3475 bp in length
10940 11039: gap of 100 bp
11040 15415: contig of 4376 bp in length
15416 15515: gap of 100 bp
15516 20495: contig of 4980 bp in length
20496 20595: gap of 100 bp
20596 25962: contig of 5367 bp in length
25963 26062: gap of 100 bp
26063 30774: contig of 4712 bp in length
30775 30874: gap of 100 bp
30875 36005: contig of 5131 bp in length
36006 36105: gap of 100 bp
36106 40573: contig of 4468 bp in length
40574 40673: gap of 100 bp
40674 47316: contig of 6643 bp in length
47317 47416: gap of 100 bp

* 47417 57551: contig of 10135 bp in length
57552 57651: gap of 100 bp
57652 67679: contig of 10028 bp in length
67709 67779: gap of 100 bp
67780 78121: contig of 10342 bp in length
78122 78221: gap of 100 bp
78222 105075: contig of 26854 bp in length
105175 105175: gap of 100 bp
124687 124687: contig of 19512 bp in length
124688 124787: gap of 100 bp
124788 137648: contig of 12861 bp in length.

FEATURES
source
1. 137648
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-2608"
/clone_lib="RPC1-11 Human Male BAC"
1. 3290
/note="assembly_fragment"
clone_end:SP6
vector_side:left
3391..5203
/note="assembly_fragment"
5304..7364
/note="assembly_fragment"
7465..10939
/note="assembly_fragment"
11040..15415
/note="assembly_fragment"
15516..20495
/note="assembly_fragment"
20596..25962
/note="assembly_fragment"
26063..30774
/note="assembly_fragment"
30875..36005
/note="assembly_fragment"
36106..40573
/note="assembly_fragment"
40674..47316
/note="assembly_fragment"
47417..57551
/note="assembly_fragment"
57652..67679
/note="assembly_fragment"
67780..78121
/note="assembly_fragment"
78222..105075
/note="assembly_fragment"
105176..124687
/note="assembly_fragment"
124788..137648
/note="assembly_fragment"
clone_end:T7
vector_side:right

ORIGIN
Query Match 2.1%; Score 28; DB 2; Length 137648;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1300 CAGATGTCACAAAAAAGAAAAAAGAAAAA 1327
|||||
DB 19515 CAGATGTCACAAAAAAGAAAAAAGAAAAA 19542
|||||
RESULT 31
AC051635/c
LOCUS AC051635 Homo sapiens chromosome 18, clone RP11-540A4, linear PRI 27-AUG-2002
DEFINITION AC051635
ACCESSION AC051635
VERSION AC051635.7 GI:17298634

KEYWORDS
SOURCE
ORGANISM

HTG.
Homo sapiens
Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 159608)

AUTHORS

Birren,B., Linton,L., Nusbaum,C. and Lander,E.

JOURNAL

Unpublished

AUTHORS

2 (bases 1 to 159608)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Gage,D.,
Dodge,S., Domano,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Hearford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Mirand,C., Mienga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE

JOURNAL

Submitted (15-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

3 (bases 1 to 159608)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,A.,
Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE

JOURNAL

Submitted (04-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

4 (bases 1 to 159608)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,A.,
Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,

Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

TITLE

JOURNAL

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

5 (bases 1 to 159608)

AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,A.,
Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

TITLE

JOURNAL

Submitted (27-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Dec 4, 2001 this sequence version replaced gi:16756307.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L5261

Center clone name: 540_A_4

FEATURES

source

Location/Qualifiers
1..159608
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone_lib="RP11-540A4"
8..106
/rpt_family="AluSg/x"
15..127
/note="<30 qual SNGL region"
75..77
/note="<30 qual SNGL region"
106..109
/note="<30 qual SNGL region"
112..228
/rpt_family="11MB2"
repeat_region
complement(563..632)
/rpt_family="MIR"

```

unsure      1053. 1057
/notes="<30 qual SNGL region"
complement(1123. 1352)
/rpt_family="MIR"
1624. 1702
/rpt_family="L2"
complement(1853. 2218)
/rpt_family="L2"
3091. 3358
/rpt_family="L1MC5"
3357. 3850
/rpt_family="L1MC1"
complement(3868. 4006)
4011. 4167
/rpt_family="L1MC7"
/rpt_family="MIR3"
complement(4235. 4560)
/rpt_family="L1R16C"
5345. 5372
/rpt_family="L1R7"
5957. 5984
/rpt_family="L1R7"
6132. 6246
/rpt_family="L1R3"
6598. 6718
/rpt_family="L1R3"
6731. 6769
/rpt_family="L1R3"
6940. 7148
/rpt_family="L1R3"
7153. 7466
/rpt_family="L1R3"
7510. 8250
/rpt_family="L1R3"
8279. 8516
/rpt_family="L1R3"
8612. 8827
/rpt_family="L1R3"
9027. 9126
/rpt_family="L1R3"
9127. 9435
/rpt_family="L1R3"
9436. 9486
/rpt_family="L1R3"
9487. 9805
/rpt_family="L1R3"
9806. 9922
/rpt_family="L1R3"
complement(9940. 9989)
/rpt_family="L1R3"

```

```

Query Match      2.1% Score 28; DB 9; Length 158608;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1300 CAGATGTCACAAAAA...AAAAAAAAA 1327
Db 82163 CAGATGTCACAAAAA...AAAAAAAAA 82136

RESULT 32
AC069032
LOCUS AC069032 187557 bp DNA linear HTG 24-AUG-2002
DEFINITION Homo sapiens chromosome 18 clone RP11-185D6 map 18, WORKING DRAFT
SEQUENCE, 44 unordered pieces.
AC069032
VERSION AC069032.4 GI:12963008
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 187557)

```

AUTHORS TITLE JOURNAL REFERENCE AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Homo sapiens chromosome 18, clone RP11-185D6
2 (bases 1 to 187557)
Unpublished
Birren,B., Linton,L., Nusbaum,C., Lander,E., Bastien,V., Bada,F.,
Anderson,S., Baldwin,J., Barna,N., Brown,A., Burkett,G.,
Boguslavskiy,L., Bouckhalter,B., Brown,A., Colangelo,M., Collins,S.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzGraham,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehotzky,J.,
Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,F.,
O'Neil,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE JOURNAL REFERENCE AUTHORS

Direct Submission
Submitted (17-MAY-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 187557)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavskiy,L., Bouckhalter,B., Brown,A.,
Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., FitzGraham,W., Gage,D., Galagan,J.,
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRoque,K., Lamazares,R., Landers,T.,
Lehotzky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

TITLE JOURNAL COMMENT

Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 18, 2001 this sequence version replaced gi:11225387.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L10383
Center clone name: 185D_6
----- Summary Statistics
Sequencing vector: M13; M77815; 2% of reads
Sequencing vector: Plasmid; n/a; 98% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 16616 bases at least Q40
Consensus quality: 176954 bases at least Q30
Consensus quality: 180848 bases at least Q20

Insert size: 167000; agarose-fp
 Insert size: 183257; sum-of-contigs
 Quality coverage: 3.7 in Q20 bases; agarose-fp
 Quality coverage: 3.4 in Q20 ba.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 44 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1294: contig of 1294 bp in length
 * 1295 1394: gap of 100 bp
 * 1395 2505: contig of 1111 bp in length
 * 2506 2606: gap of 100 bp
 * 2606 3992: contig of 1387 bp in length
 * 3993 4092: gap of 100 bp
 * 4093 5527: contig of 1435 bp in length
 * 5528 5627: gap of 100 bp
 * 5628 6675: contig of 1048 bp in length
 * 6676 6775: gap of 100 bp
 * 6776 8026: contig of 1251 bp in length
 * 8027 8126: gap of 100 bp
 * 8127 9162: contig of 1036 bp in length
 * 9163 9262: gap of 100 bp
 * 9263 10351: contig of 1089 bp in length
 * 10352 10451: gap of 100 bp
 * 10452 11757: contig of 1306 bp in length
 * 11758 11857: gap of 100 bp
 * 11858 12978: contig of 1021 bp in length
 * 12979 12979: gap of 100 bp
 * 12979 14159: contig of 1181 bp in length
 * 14160 14259: gap of 100 bp
 * 14260 15818: contig of 1559 bp in length
 * 15819 15918: gap of 100 bp
 * 15919 17145: contig of 1227 bp in length
 * 17146 17245: gap of 100 bp
 * 17246 18785: contig of 1520 bp in length
 * 18786 18865: gap of 100 bp
 * 18866 20138: contig of 1273 bp in length
 * 20139 20238: gap of 100 bp
 * 20239 21879: contig of 1641 bp in length
 * 21880 21880: gap of 100 bp
 * 21880 23877: contig of 1898 bp in length
 * 23878 23977: gap of 100 bp
 * 23978 25680: contig of 1703 bp in length
 * 25681 25780: gap of 100 bp
 * 25781 28099: contig of 2319 bp in length
 * 28100 28199: gap of 100 bp
 * 28200 30797: contig of 2598 bp in length
 * 30798 30897: gap of 100 bp
 * 30898 33191: contig of 2294 bp in length
 * 33192 33191: gap of 100 bp
 * 33292 36520: contig of 3229 bp in length
 * 36521 36620: gap of 100 bp
 * 36621 38842: contig of 2222 bp in length
 * 38843 38943: gap of 100 bp
 * 38943 42470: contig of 3528 bp in length
 * 42471 42570: gap of 100 bp
 * 42571 45581: contig of 3011 bp in length
 * 45582 45681: gap of 100 bp
 * 45682 48005: contig of 2324 bp in length
 * 48006 48105: gap of 100 bp
 * 48106 51907: contig of 3802 bp in length
 * 51908 52007: gap of 100 bp
 * 52008 55311: contig of 3304 bp in length
 * 55312 55411: gap of 100 bp
 * 55412 59353: contig of 3942 bp in length
 * 59354 59454: gap of 100 bp
 * 59454 63507: contig of 4053 bp in length
 * 63507 63606: gap of 100 bp
 * 63607 68052: contig of 4446 bp in length

* 68053 68152: gap of 100 bp
 * 68153 72726: contig of 4574 bp in length
 * 72727 72826: gap of 100 bp
 * 72827 78593: contig of 5767 bp in length
 * 78594 78694: gap of 100 bp
 * 78694 84391: contig of 5698 bp in length
 * 84392 84491: gap of 100 bp
 * 84492 92490: contig of 7959 bp in length
 * 92491 92590: gap of 100 bp
 * 92591 99730: contig of 7140 bp in length
 * 99731 99830: gap of 100 bp
 * 99831 107602: contig of 7772 bp in length
 * 107603 107703: gap of 100 bp
 * 107703 116403: contig of 8701 bp in length
 * 116404 126259: contig of 9756 bp in length
 * 126260 126359: gap of 100 bp
 * 126360 136166: contig of 9807 bp in length
 * 136167 136266: gap of 100 bp
 * 136267 145578: contig of 9312 bp in length
 * 145579 145678: gap of 100 bp
 * 145679 157186: contig of 11508 bp in length
 * 157187 157286: gap of 100 bp
 * 157287 170785: contig of 13499 bp in length
 * 170786 170885: gap of 100 bp
 * 170886 187557: contig of 16672 bp in length.

FEATURES

source
 1..187557
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="18"
 /map="18"
 /clone="RP11-185D6"
 /clone_lib="RPC1-11 Human Male BAC"
 1..1294
 /note="assembly_fragment"
 1395..2505
 /note="assembly_fragment"
 2606..3992
 /note="assembly_fragment"
 4093..5527
 /note="assembly_fragment"

Query Match 2.1%; Score 28; DB 2; Length 187557;
 Best Local Similarity 100.0%; Pred. No. 0.00014;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1300 CAGATGTCACAAAAA 1327
 |||||
 Db 28971 CAGATGTCACAAAAA 28998

RESULT 33
 AC068678/c 193978 bp DNA linear HTG 31-OCT-2001
 LOCUS Homo sapiens chromosome 18 clone RP11-104N14 map 18, WORKING DRAFT
 DEFINITION SEQUENCE, 5 unordered pieces.
 AC068678
 VERSION AC068678.3 GI:16506881
 KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 193978)
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 TITLE Homo sapiens chromosome 18, clone RP11-104N14
 JOURNAL Unpublished
 2 (bases 1 to 193978)
 REFERENCE Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
 Boguslavskiy, L., Boukhalter, B., Brown, A., Burkett, G.,

Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., LaRoque, K., Lanazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneus, J., Mihova, T., Miranda, C., Mienga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (06-MAY-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Oct 28, 2001 this sequence version replaced gi:13249473.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIER

Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L1295

Center clone name: 104 N.14

----- Summary Statistics

Sequencing vector: M13; M7815; 4% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 191018 bases at least Q40
 Consensus quality: 192308 bases at least Q30
 Consensus quality: 192875 bases at least Q20
 Insert size: 190000; agarose-fp
 Insert size: 193578; sum-of-ctdigs
 Quality coverage: 7.3 in Q20 bases; agarose-fp
 Quality coverage: 7.1 in Q20 ba

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 5 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 59978: contig of 59978 bp in length
 59979 60078: gap of 100 bp
 60079 83937: contig of 23859 bp in length
 83938 84037: gap of 100 bp
 84038 113621: contig of 29584 bp in length
 113622 113721: gap of 100 bp
 113722 147302: contig of 33581 bp in length
 147303 147402: gap of 100 bp
 147403 193978: contig of 46576 bp in length.

FEATURES

source

1. 193978
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="18"
 /map="18"
 /clone="RP11-104N14"
 /clone_lib="RPC1-11 Human Male BAC"
 1. 59978
 /note="assembly_fragment"

misc_feature

clone end:SP6
 vector_side:left
 60079..83937
 /note="assembly_fragment"
 84038..113621
 /note="assembly_fragment"
 113722..147302
 /note="assembly_fragment"
 147403..193978
 /note="assembly_fragment"
 clone end:T7
 vector_side:right"

ORIGIN

Query Match 2.1%; Score 28; DB 2; Length 193978;
 Best Local Similarity 100.0%; Pred. No. 0.00014;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1300 CAGATGTCACAAAAAAGAAAAAAGAAAAA 1327

Db 20268 CAGATGTCACAAAAAAGAAAAAAGAAAAA 20241

RESULT 34

AC011814
 LOCUS AC011814 209462 bp DNA linear HTG 26-MAY-2000
 DEFINITION Homo sapiens chromosome 18 clone RP11-153D16 map 18, WORKING DRAFT
 SEQUENCE, 40 unordered pieces.

AC011814

AC011814.3 GI:8072439

VERSION HTG; HTGS PHASE1; HTGS_DRAFT.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Theria; Primates; Catarrhini; Hominoidea; Homo.

Barren, B., Linton, L., Nusbaum, C. and Lander, E.

Unpublished

2 (bases 1 to 209462)

Barren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,

Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Boukhgalter, B.,

Brown, A., Castle, A., Colangelo, M., Collins, S., Collamore, A.,

Cooke, P., DeArellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,

Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D.,

Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,

Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,

Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,

McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J.,

Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,

Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,

Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,

Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,

Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (15-OCT-1999) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On May 25, 2000 this sequence version replaced gi:7321509.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIER

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L1307

Center clone name: 153 D.16

----- Summary Statistics

Sequencing vector: M13; M7815; 97% of reads

Sequencing vector: Plasmid; n/a; 8-0.1% of reads

3.29835082458771Chemistry: Dye-primer-amersham; 8% of reads

Chemistry: Dye-terminator Big Dye; 92% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 172516 bases at least Q40

Consensus quality: 188069 bases at least Q30

Consensus quality: 195901 bases at least Q20

Insert size: 175000; agarose-fp

Insert size: 205562; sum-of-contigs

Quality covera.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 40 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

```

1 1353: contig of 1353 bp in length
1354 1453: gap of 100 bp
1454 2778: contig of 1325 bp in length
2779 2878: gap of 100 bp
2879 4211: contig of 1333 bp in length
4212 4311: gap of 100 bp
4312 5622: contig of 1311 bp in length
5623 5722: gap of 100 bp
5723 7256: contig of 1534 bp in length
7257 7356: gap of 100 bp
7357 8828: contig of 1472 bp in length
8829 10325: contig of 1297 bp in length
10326 10325: gap of 100 bp
10326 11376: contig of 1051 bp in length
11377 12644: gap of 100 bp
12644 12743: gap of 100 bp
12744 13857: contig of 1114 bp in length
13858 13957: gap of 100 bp
13958 15355: contig of 1398 bp in length
15356 15455: gap of 100 bp
15456 17127: contig of 1672 bp in length
17128 17227: gap of 100 bp
17228 19117: contig of 1890 bp in length
19118 19217: gap of 100 bp
19218 20984: contig of 1767 bp in length
20985 21084: gap of 100 bp
21085 23101: contig of 2017 bp in length
23102 23201: gap of 100 bp
23202 25331: contig of 2130 bp in length
25332 25431: gap of 100 bp
25432 27214: contig of 1783 bp in length
27215 27314: gap of 100 bp
27315 29160: contig of 1846 bp in length
29161 29260: gap of 100 bp
29261 32570: contig of 3310 bp in length
32571 32670: gap of 100 bp
32671 34880: contig of 2210 bp in length
34881 34981: gap of 100 bp
34982 38325: contig of 3345 bp in length
38326 38425: gap of 100 bp
38426 42527: contig of 4102 bp in length
42528 42627: gap of 100 bp
42628 46099: contig of 3472 bp in length
46100 46199: gap of 100 bp
46200 49618: contig of 3419 bp in length
49619 49718: gap of 100 bp
49719 53310: contig of 3592 bp in length
53311 53410: gap of 100 bp
53411 56772: contig of 3362 bp in length
56773 56872: gap of 100 bp
56873 60358: contig of 3486 bp in length
60359 60458: gap of 100 bp
60459 66617: contig of 6159 bp in length
66618 66717: gap of 100 bp
66718 75403: contig of 8686 bp in length

```

```

* 75404 75503: gap of 100 bp
* 75504 83115: contig of 7612 bp in length
* 83116 83215: gap of 100 bp
* 83216 91822: contig of 8607 bp in length
* 91823 91922: gap of 100 bp
* 91923 101265: contig of 9343 bp in length
* 101266 101365: gap of 100 bp
* 101366 111447: contig of 10082 bp in length
* 111448 111547: gap of 100 bp
* 111548 121183: contig of 9636 bp in length
* 121184 121283: gap of 100 bp
* 121284 133655: contig of 12372 bp in length
* 133656 133755: gap of 100 bp
* 133756 147893: contig of 14138 bp in length
* 147894 147993: gap of 100 bp
* 147994 160303: contig of 12309 bp in length
* 160304 174936: contig of 14534 bp in length
* 174937 175036: gap of 100 bp
* 175037 189972: contig of 14936 bp in length
* 189973 209462: contig of 19390 bp in length.
* 190073 209462: contig of 19390 bp in length.
FEATURES
      source
      1..209462
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /chromosome="18"
      /map="18"
      /clone="RP11-153D16"
      /clone_lib="RPC1-11 Human Male BAC"
      1..1353
      /note="assembly_fragment"
      1454..2778
      /note="assembly_fragment"
      2879..4211
      /note="assembly_fragment"
      4312..5622
      /note="assembly_fragment"
      5723..7256
      /note="assembly_fragment"
      7357..8828
      /note="assembly_fragment"
      8929..10225
      /note="assembly_fragment"
      10326..11376
      /note="assembly_fragment"
      11477..12643
      /note="assembly_fragment"
      12744..13857
      /note="assembly_fragment"
      13958..15355
      /note="assembly_fragment"
      15456..17127
      /note="assembly_fragment"
      17228..19117
      /note="assembly_fragment"
      19218..20984
      /note="assembly_fragment"
      21085..23101
      /note="assembly_fragment"
      23202..25331
      /note="assembly_fragment"
      25432..27214
      /note="assembly_fragment"
      27315..29160
      /note="assembly_fragment"
      29261..32570
      /note="assembly_fragment"
      32671..34880
      /note="assembly_fragment"
      34981..38325
      /note="assembly_fragment"

```

```

misc_feature 38426..42527
              /note="assembly_fragment"
misc_feature 42628..46099
              /note="assembly_fragment"
misc_feature 46200..49618
              /note="assembly_fragment"

Query Match      2.1%; Score 28; DB 2; Length 209462;
Best Local Similarity 100.0%; Pred. No. 0.00014; Indels 0; Gaps 0;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1300 CAGATGTCAAAAAAAAAAAAAAAAAAAAA 1327
      |||||
Db 54971 CAGATGTCAAAAAAAAAAAAAAAAAAAAA 54998

RESULT 35
BC060625 Mus musculus Tnf receptor-associated factor 2, mRNA ROD 12-NOV-2003
LOCUS IMAGE:6831932, partial cds.
ACCESSION BC060625
VERSION BC060625.1 GI:38197727
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 4150)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marasina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,I.B.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raja,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Hale,S.G., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S.G., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalil,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
PUBMED
2 (bases 1 to 4150)
Strausberg,R.
Direct Submission
Submitted (31-OCT-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapsb-x@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
Thomas L. Casavant.
Web site: http://genome.uiowa.edu
Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
Bonaldo,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A.,
Fisher,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K.,
Scheetz,T., Smith,C., Snir,E., Tack,D., Trout,K., Walters,J.,
Casavant,T., Soares,M.B.

misc_feature 38426..42527
              /note="assembly_fragment"
misc_feature 42628..46099
              /note="assembly_fragment"
misc_feature 46200..49618
              /note="assembly_fragment"

Query Match      2.0%; Score 27; DB 10; Length 4150;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1301 AGATGTCAAAAAAAAAAAAAAAAAAAAA 1327
      |||||
Db 4111 AGATGTCAAAAAAAAAAAAAAAAAAAAA 4137

RESULT 36
HSM803417 Homo sapiens mRNA; cDNA DKFZp313H168 (from clone DKFZp313H168).
LOCUS HSM803417 4685 bp mRNA linear PRI 13-MAY-2003
DEFINITION Homo sapiens mRNA; cDNA DKFZp313H168 (from clone DKFZp313H168).
ACCESSION AL832110
VERSION AL832110.1 GI:21732653
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4685)
Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
Direct Submission
Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp313H168) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.

FEATURES
source 1..4685
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="DKFZp313H168"
              /tissue_type="human bone marrow"
              /clone_lib="313 (synonym: blcc2). Vector pTriplEx2; host
              DH10B; sites SfiIA + SfiIB"
              /dev_stage="adult"
              /dev_stage="adult"
polyA signal 4647..4652
polyA_site 4662
ORIGIN
Query Match      2.0%; Score 27; DB 9; Length 4685;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>
Series: Plate: Row: Column: 0
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

```

FEATURES
source 1..4150
              /organism="Mus musculus"
              /mol_type="mRNA"
              /strain="C57BL/6"
              /db_xref="taxon:10090"
              /clone="IMAGE:6831932"
              /tissue_type="Brain, mouse, 13.5,14.5,16.5,17.5 dpc"
              /clone_lib="NIH_BMAP_FY0"
              /lab_host="DH10B"
              /note="Vector: pYX-ASC"
Location/Qualifiers

```

```

ORIGIN
Query Match      2.0%; Score 27; DB 10; Length 4150;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1301 AGATGTCAAAAAAAAAAAAAAAAAAAAA 1327
      |||||
Db 4111 AGATGTCAAAAAAAAAAAAAAAAAAAAA 4137

```

```

RESULT 36
HSM803417 Homo sapiens mRNA; cDNA DKFZp313H168 (from clone DKFZp313H168).
LOCUS HSM803417 4685 bp mRNA linear PRI 13-MAY-2003
DEFINITION Homo sapiens mRNA; cDNA DKFZp313H168 (from clone DKFZp313H168).
ACCESSION AL832110
VERSION AL832110.1 GI:21732653
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4685)
Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
Fobo,G., Han,M. and Wiemann,S.
Direct Submission
Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp313H168) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.

```

```

FEATURES
source 1..4685
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="DKFZp313H168"
              /tissue_type="human bone marrow"
              /clone_lib="313 (synonym: blcc2). Vector pTriplEx2; host
              DH10B; sites SfiIA + SfiIB"
              /dev_stage="adult"
              /dev_stage="adult"
polyA signal 4647..4652
polyA_site 4662
ORIGIN
Query Match      2.0%; Score 27; DB 9; Length 4685;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

repeat_region complement(10418..10545)
                /evidence=not_experimental
                /rpt_family="MER58C"
repeat_region 10546..10585
                /evidence=not_experimental
                /rpt_family="Alu"
repeat_region 10708..10797
                /evidence=not_experimental
                /rpt_family="L2"
repeat_region complement(10837..11091)
                /evidence=not_experimental
                /rpt_family="AluSc"
repeat_region 11771..11813
                /evidence=not_experimental
                /rpt_family="AT_rich"
repeat_region 12272..12292
                /evidence=not_experimental
                /rpt_family="AT_rich"
repeat_region complement(12629..12706)
                /evidence=not_experimental
                /rpt_family="MIR"
repeat_region 12903..13167
                /evidence=not_experimental
                /rpt_family="AluJo"
repeat_region 13173..13453
                /evidence=not_experimental
                /rpt_family="AluSg"
repeat_region 13871..13920
                /evidence=not_experimental
                /rpt_family="AT_rich"
repeat_region complement(14938..15193)
                /evidence=not_experimental
                /rpt_family="AluSg"
repeat_region 15379..15453
                /evidence=not_experimental
                /rpt_family="L2"
repeat_region complement(15501..15754)
                /evidence=not_experimental
                /rpt_family="L2"
repeat_region complement(16177..16924)
                /evidence=not_experimental
                /rpt_family="MER31-internal"
repeat_region complement(17232..17377)
                /evidence=not_experimental
                /rpt_family="MER31-internal"
repeat_region 18854..18955
                /evidence=not_experimental
                /rpt_family="AT_rich"
repeat_region 19314..19624
                /evidence=not_experimental
                /rpt_family="AluSx"
repeat_region 20973..21071
                /evidence=not_experimental
                /rpt_family="MIR"
repeat_region complement(22061..22369)
                /evidence=not_experimental
                /rpt_family="AluSx"
repeat_region 22428..22463
                /evidence=not_experimental
                /rpt_family="AT_rich"
repeat_region complement(22469..22781)
                /evidence=not_experimental
                /rpt_family="AluJo"
repeat_region complement(22820..23154)
                /evidence=not_experimental
                /rpt_family="MER7A"
repeat_region 23362..23397
                /evidence=not_experimental
                /rpt_family="AT_rich"
repeat_region 23468..23503
                /evidence=not_experimental
                /rpt_family="AT_rich"
repeat_region 24148..24175

```

```

repeat_region /evidence=not_experimental
                /rpt_family="Alu"
                complement(24576..24879)
                /evidence=not_experimental
                /rpt_family="AluSx"
                24980..25344
                /evidence=not_experimental
                /rpt_family="MLT2B"
                complement(25365..25422)
                /evidence=not_experimental
                /rpt_family="MIR"
                complement(25372..25487)
                /evidence=not_experimental
                /rpt_family="L2"
                25515..25637
                /evidence=not_experimental
                /rpt_family="MER94"
                25822..25868
                /evidence=not_experimental
                /rpt_family="Alu"
                26514..26818
                /evidence=not_experimental
                /rpt_family="AluYas"
                27318..27344
                /evidence=not_experimental
                /rpt_family="(TTTG)n"
                27983..28027
                /evidence=not_experimental
                /rpt_family="AT_rich"
                complement(28187..28382)
                /evidence=not_experimental
                /rpt_family="MER20"
                28857..28993
                /evidence=not_experimental
                /rpt_family="FLAM_C"
                29511..29565
                /evidence=not_experimental
                /rpt_family="(TG)n"
                29582..29620
                /evidence=not_experimental
                /rpt_family="AT_rich"
                29781..29923
                /evidence=not_experimental
                /rpt_family="MIR"
                complement(30500..30659)
                /evidence=not_experimental
                /rpt_family="MIR"
                complement(30851..31133)
                /evidence=not_experimental
                /rpt_family="THE1C"
                31305..31348
                /evidence=not_experimental
                /rpt_family="AT_rich"
                32634..32725

```

Query Match 2.0%; Score 27; DB 9; Length 105199;
 Best Local Similarity 100.0%; Pred. No. 0.00053;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1301 AGATGTCACAAAAA 1327
 Db 96086 AGATGTCACAAAAA 96112

RESULT 39

AC112138
 LOCUS AC112138 133897 bp DNA linear PRI 29-MAY-2002
 DEFINITION Homo sapiens 3 BAC RP11-715D1 (Roswell Park Cancer Institute Human
 BAC Library) complete sequence.
 ACCESSION AC112138
 VERSION AC112138.3 GI:21240534
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS

1 (bases 1 to 133897)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alzubrooks,S.L., Amaratunga,H.C., Are,J.R., Ayele,M., Banks,T., Barbakid,J., Benton,J., Bimage,K., Blumenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieleva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.P., Carter,M., Cavaros,S.R., Chacko,V., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Kensi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W., Loulsegad,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,M., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,I., Peters,L., Pickens,R., Primus,E., Pu,F.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojupokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstock,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 133897)
Worley,K.C.
Direct Submission
Submitted (19-FEB-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 133897)
Worley,K.C.
Direct Submission
Submitted (09-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 133897)
Worley,K.C.
Direct Submission
Submitted (29-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 29, 2002 this sequence version replaced gi:20127825.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones.

Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished), for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at UKR:

<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.

FEATURES	Location/Qualifiers
source	1..133897
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="3"
	/clone="RP11-715D1"
repeat_region	2057..2229
	/rpt_family="MIR"
repeat_region	2318..2351
	/rpt_family="(CA)n"
repeat_region	2772..3078
	/rpt_family="AluY"
repeat_region	3294..3585
	/rpt_family="AluSx"
repeat_region	5345..5474
	/rpt_family="L1MC5"
repeat_region	6402..6494
	/rpt_family="MIR"
repeat_region	6906..6929
	/rpt_family="AT-rich"
repeat_region	complement(7169..7337)
	/rpt_family="AluSq"
repeat_region	7412..7457
	/rpt_family="WADE1"
repeat_region	7459..8190
	/rpt_family="L2"
repeat_region	8730..9295
	/rpt_family="L2"
repeat_region	9300..9384
	/rpt_family="(TA)n"
repeat_region	9440..9477
	/rpt_family="(CA)n"
repeat_region	10044..10137
	/rpt_family="AluY"
repeat_region	12868..12894
	/rpt_family="AT-rich"
repeat_region	13556..13770
	/rpt_family="L2"
repeat_region	13948..13981
	/rpt_family="(TG)n"
repeat_region	13982..14012

[illegible]

6	<800	6382	6519	2067	2160	4060	4080	934	979	10740	10687
2332	2547	512	<800	8527	8714	6	<800	6627	6519	2316	2399
3139	3178	449	<800	1375	1342	3386	3376	1468	1442	3583	3433
4975	5077	1133	1128	893	930	675	<800	2168	2121	2782	2789
2418	2547	2746	2762	4253	4526	707	<800	951	979	188	<800
1931	1905	2037	2121	436	<800	1393	1380	2339	2439	319	<800
7381	7438	2845	2869	925	930	8957	8700	2842	2869		
1269	1212	3682	3728	2899	2789	3442	3662	3748	3728		
3691	3662	3956	3889	16122	16215	4672	4584	1164	1128		
3330	3376	1008	979	2045	2160	2212	2288	3610	3585		
8174	8159	5364	5343	3498	3433	4156	4080	3583	3585		
5231	5077	2323	2348	650	<800	1444	1474	457	<800		
2843	2906	1357	1309	13741	13713			7339	7196		
5157	5077	5192	5120	60	<800			559	<800		
181	<800	10560	10611	5778	6013			1883	1917		
1636	1608	758	<800	2421	2399			8831	8854		
556	<800	10106	10165	1491	1472			1515	1442		
1494	1474	5709	5695	2486	2399						
360	<800	342	<800	1861	1904						
5006	5077	2678	2762	12854	12717						
638	<800	1893	1917	16	<800						
5012	5077	304	<800	18	<800						
256	<800	3190	3256	1286	1265						
6568	6477	2488	2541	3630	3895						
414	<800	5865	5695	1037	1057						
2045	2079	1457	1442	4025	4293						
17911	18009	5628	5695	3747	3895						
756	<800	496	<800	6587	6827						
5394	5422	886	979	2148	2160						
1063	1053	4473	4390	10135	10124						
1504	1474	1171	1128	1908	1904						
3606	3376	465	<800	719	<800						
4911	5077	7137	7196	2175	2160						
208	<800	1039	979	2907	2789						
5151	5077	234	<800	1064	1057						
918	923	4265	4216	3648	3895						

FEATURES
source

Location/Qualifiers
1..152772
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"

Query Match 2.0%; Score 27; DB 9; Length 152772;
Best Local Similarity 100.0%; Pred. No. 0.00055;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1301 AGATGTCACAAAAA 1327
Db 146084 AGATGTCACAAAAA 146058

RESULT 41

AC063923/c
LOCUS AC063923 153448 bp DNA linear PRI 03-SEP-2002
DEFINITION Homo sapiens 3 BAC RP11-59B19 (Roswell Park Cancer Institute Human BAC Library) complete sequence.
ACCESSION AC063923
VERSION AC063923.21 GI:19807681
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 153448)
AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-ouman, F.R., Allen, C., Alsbrooks, S.L., Amaratunga, H.C., Are, J.R., Ayele, M., Banks, T., Barbaria, J., Benton, J., Binage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carton, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyie, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,

Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, P., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., He, X., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseghe, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, M., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Moore, S., Morgan, M., Moorish, T., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okuwonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojebokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I., Sodergren, E., Sonaika, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Naylor, S.L., Weinstein, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 153448)
Worley, K.C.

Direct Submission
Submitted (22-APR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 153448)
Worley, K.C.

Direct Submission
Submitted (26-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 153448)
Worley, K.C.

Direct Submission
Submitted (29-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 153448)
Worley, K.C.

Direct Submission
Submitted (03-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Mar 29, 2002 this sequence version replaced gi:19718593.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,

unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT	
FEATURES	Location/Qualifiers
source	1..153448
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosomes="3"
	/clone="RP11-59E19"
misc_feature	1..2019
	/note="overlaps bases 73308..75323 of clone AC076971"
	/function="clone overlap"
repeat_region	complement(133..416)
	/rpt_family="AluJb"
repeat_region	419..471
	/rpt_family="L1MCS"
repeat_region	complement(472..633)
	/rpt_family="FRAM"
repeat_region	634..718
	/rpt_family="L1MCS"
repeat_region	780..1157
	/rpt_family="MLT1D"
repeat_region	complement(1434..2333)
	/rpt_family="L1M4"
repeat_region	complement(2402..2577)
	/rpt_family="L1M4"
repeat_region	2583..2624
	/rpt_family="(TA)n"
repeat_region	complement(2625..2872)
	/rpt_family="AluSx"
repeat_region	2873..2900
	/rpt_family="(TA)n"
repeat_region	complement(2961..3085)
	/rpt_family="L1M4"
repeat_region	complement(3182..3481)
	/rpt_family="AluJo"
repeat_region	complement(4166..4456)
	/rpt_family="AluSx"
repeat_region	4686..4865
	/rpt_family="(TTCTCC)n"
repeat_region	complement(5577..5875)
	/rpt_family="AluY"
repeat_region	complement(5879..6244)
	/rpt_family="MLT1E"
repeat_region	complement(6339..6432)
	/rpt_family="MLT1E"
repeat_region	6581..7060
	/rpt_family="ITR40a"
repeat_region	7537..7566
	/rpt_family="(TTTTTA)n"
repeat_region	complement(7567..7714)
	/rpt_family="AluSx"

REFERENCE 4 (bases 1 to 174531)

[illegible]

462	<800	3436	3418	449	<
-----	------	------	------	-----	---

[illegible]

3573	3730	314	<800	1586	1577	2806
3957	4016	611	<800	6006	6016	2746
231	<800	1206	1185	2640	2612	2076
5379	5264	11764	11362	989	1000	2806
208	<800	645	<800	2203	2351	3860
541	<800	2775	2642	6314	6297	4049
3765	3910	533	<800	5814	5690	1000
4033	4185	4992	4944	773	<800	5318
604	<800	3290	3297	2358	2351	2351
541	<800	2896	2958	3526	3509	1329
6856	6776	6623	6599	1475	1475	5136
9840	9668	1188	1185	5668	5690	10559
479	<800	2656	2642	4882	4809	795
1716	1678	529	<800	2728	2806	10067
6109	6092	10269	10030	998	1000	
913	919	186	<800	2264	2351	
4352	4330	5513	5517	7957	7979	
1536	1530	2479	2486	3668	3860	
2338	2345	8122	8148	2347	2351	
16968	17243	5075	5133	2766	2806	
9	<800	2032	2025	1778	1790	
624	<800	747	765	1202	1180	
5148	5070	845	868	7646	7689	
564	<800	6360	6444	6107	6016	
2737	2743	486	<800	2081	2076	
6213	6092	16691	16973	61	<800	
1102	1115	1625	1613	223	<800	
8039	8044	4695	4944	4130	4049	
688	<800	2848	2844	274	<800	
4877	4841	2532	2642	306	<800	
1671	1678	3140	3156	4169	4049	
1685	1678	4975	4944	3560	3655	
976	919	2418	2486	9213	9229	
4328	4330	1931	1908	801	822	
1376	1335	7391	7416	2192	2351	
893	919	1269	1185	1617	1577	
4253	4330	3691	3700	1191	1180	

Query Match

Best Local Similarity

Matches

27;

Conservative

0;

Mismatches

0;

Indels

0;

Gaps

0;

QY

1301

AGATGTC

AAAAAAAAAAAAAAAAAAAAA

1327

Db

55262

AGATGTC

AAAAAAAAAAAAAAAAAAAAA

55236

RESULT

43

AC010959

LOCUS

AC010959

DEFINITION

Homo sapiens clone RP11-7D10, WORKING DRAFT SEQUENCE, 25 unordered pieces.

AC010959

AC010959.3

GI:7209492

HTG; HTGS PHASE1; HTGS_DRAFT.

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

(bases 1 to 177672)

1

Homo sapiens, clone RP11-7D10

Unpublished

2

(bases 1 to 177672)

1

Bairren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Bouckgalter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,B., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission

Submitted (28-SEP-1999); Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Mar 8, 2000 this sequence version replaced gi:6978150.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

TITLE

JOURNAL

COMMENT

FEATURES	source	Location/Qualifiers
		1. 177672
		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
		/clone="RP11-7D10"
		/clone_lib="RPC1-11 Human Male BAC"
		1. 1129
misc_feature		/note="assembly_fragment"
misc_feature		1230..2365
misc_feature		/note="assembly_fragment"
misc_feature		2466..4110
misc_feature		/note="assembly_fragment"
misc_feature		4211..5970
misc_feature		/note="assembly_fragment"
misc_feature		6071..7611
misc_feature		/note="assembly_fragment"
misc_feature		7712..9717
misc_feature		/note="assembly_fragment"
misc_feature		9818..12688
misc_feature		/note="assembly_fragment"
misc_feature		12789..16615
misc_feature		/note="assembly_fragment"
misc_feature		16716..20067
misc_feature		/note="assembly_fragment"
misc_feature		20168..23688
misc_feature		/note="assembly_fragment"
misc_feature		23789..27981
misc_feature		/note="assembly_fragment"
misc_feature		28082..31601
misc_feature		/note="assembly_fragment"
misc_feature		31702..37510
misc_feature		/note="assembly_fragment"
misc_feature		37611..42701
misc_feature		/note="assembly_fragment"
misc_feature		42802..50137
misc_feature		/note="assembly_fragment"
misc_feature		50238..57522
misc_feature		/note="assembly_fragment"
misc_feature		57623..66946
misc_feature		/note="assembly_fragment"
misc_feature		67047..74537
misc_feature		/note="assembly_fragment"
misc_feature		74638..78928
misc_feature		/note="assembly_fragment"
misc_feature		clone_end:SP6
misc_feature		vector_side:left"
misc_feature		79029..89116
misc_feature		/note="assembly_fragment"
misc_feature		89217..98944
misc_feature		/note="assembly_fragment"
misc_feature		99045..113038
misc_feature		/note="assembly_fragment"
misc_feature		113139..131560
misc_feature		/note="assembly_fragment"
misc_feature		131661..150982
misc_feature		/note="assembly_fragment"
misc_feature		clone_end:T7
misc_feature		vector_side:right"
misc_feature		151083..177672
misc_feature		/note="assembly_fragment"
		ORIGIN
		Query Match 2.0%; Score 27; DB 2; Length 17
		Best Local Similarity 100.0%; Pred. No. 0.00055;
		Matches 27; Conservative 0; Mismatches 0; Indels
		1301 AGATGTCMAAAAAAAAAAAAAAAAAAAAA 1327

```

Db 102188 AGATGTCACAAAAA 102214
|||||
REFERENCE 1
AUTHORS Heinrich,G. and Kerb,R.
TITLE Methods for the treatment of cancer with irinotecan based on CYP3A5
JOURNAL Patent: WO 03013534-A 669 20-FEB-2003;
FEATURES Epidauros Biotechnologie AG (DE)
SOURCE Location/Qualifiers
1. .189013
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 2.0%; Score 27; DB 6; Length 189013;
Best Local Similarity 100.0%; Pred. No. 0.00056;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1301 AGATGTCACAAAAA 1327
|||||
Db 123937 AGATGTCACAAAAA 123963
|||||
RESULT 46
AX707901 189013 bp DNA linear PAT 04-APR-2003
LOCUS Sequence 669 from Patent WO03013536.
DEFINITION AX707901
ACCESSION AX707901
VERSION AX707901.1 GI:29563976
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Heinrich,G. and Kerb,R.
TITLE Methods for treatment of cancer using irinotecan based on UGT1A1
JOURNAL Patent: WO 03013536-A 669 20-FEB-2003;
FEATURES Epidauros Biotechnologie AG (DE)
SOURCE Location/Qualifiers
1. .189013
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 2.0%; Score 27; DB 6; Length 189013;
Best Local Similarity 100.0%; Pred. No. 0.00056;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1301 AGATGTCACAAAAA 1327
|||||
Db 123937 AGATGTCACAAAAA 123963
|||||
RESULT 47
AC025277 189032 bp DNA linear PRI 25-NOV-2003
LOCUS Homo sapiens chromosome 16 clone CTD-3229J4, complete sequence.
DEFINITION AC025277
ACCESSION AC025277
VERSION AC025277.10 GI:38502330
KEYWORDS HTG
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS DOE Joint Genome Institute, Stanford Human Genome Center and Los
Alamos National Laboratory.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (08-MAR-2000) Production Sequencing Facility, DOE Joint

```

```

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 189032)
DOE Joint Genome Institute.
Direct Submission
Submitted (20-FEB-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 189032)
DOE Joint Genome Institute.
Direct Submission
Submitted (06-MAR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
5 (bases 1 to 189032)
DOE Joint Genome Institute.
Direct Submission
Submitted (05-SEP-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
6 (bases 1 to 189032)
Stanford Human Genome Center and Los Alamos National Laboratory.
Direct Submission
DOE Joint Genome Institute
Submitted (25-NOV-2003) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Nov 25, 2003 this sequence version replaced gi:22725948.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center and Los Alamos
National Laboratory
www.shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.
Location/Qualifiers
1. 189032
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"
/clones="CID-322974"

FEATURES
source
Query Match 2.0%; Score 27; DB 9; Length 189032;
Best Local Similarity 100.0%; Pred. No. 0.00056;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1301 AGATGTCACAAAAA 1327
|||||
Db 123946 AGATGTCACAAAAA 123972
|||||

RESULT 48
AC110206/c
LOCUS 199045 bp DNA linear HTG 17-DEC-2003
DEFINITION Mus musculus chromosome 19 clone RP23-155H17 map 19, *** SEQUENCING
IN PROGRESS ***, 2 ordered pieces.
ACCESSION AC110206
VERSION AC110206.18 GI:39841427
KEYWORDS HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 199045)
Mus musculus chromosome 19, clone RP23-155H17
Unpublished
2 (bases 1 to 199045)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavskiy, L., Bouckhgalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cook, P., DeArrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hulme, W., Illiev, I., Johnson, R., Jones, C.,
Ginde, S., Gord, S., Coyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Illiev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lander, E., Abouelleil, A., Allen, N.,
Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C.,
Macedonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, P., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zaitoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (10-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 199045)
Birren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavskiy, L., Bouckhgalter, B., Camarata, J., Chang, J., Choepel, Y.,
Collymore, A., Cook, A., Cooke, P., Corum, B., DeArrellano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Illiev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., MacLean, C.,
Macedonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
Wyman, D., Young, G., Zaitoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (17-DEC-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 15, 2003 this sequence version replaced gi:37202218.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L20741
Center clone name: 155_H_17
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 136971: contig of 136971 bp in length
* 136972 137071: gap of 100 bp
* 137072 199045: contig of 61974 bp in length.
Location/Qualifiers
1. 199045
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="19"
/map="19"
FEATURES
source

```


Fri Sep 17 10:18:23 2004

repeat_region	18556..18866	/rpt_family="B4"	1301	AGATGTCACAAAAA	1327
repeat_region	20123..20175	/rpt_family="5S"	Db	197260	AGATGTCACAAAAA
repeat_region	20397..20697	/rpt_family="MER1_type"			
repeat_region	20899..21004	/rpt_family="MER1_type"			
repeat_region	21093..21197	/rpt_family="MER2_type"			
repeat_region	22734..23067	/rpt_family="RMR12"			
repeat_region	23155..23221	/rpt_family="RMR12"			
repeat_region	25357..25466	/rpt_family="Alu"			
repeat_region	25637..25787	/rpt_family="MER1_type"			
repeat_region	28526..28927	/rpt_family="L1"			
repeat_region	29467..29563	/rpt_family="Alu"			
repeat_region	31412..31561	/rpt_family="Alu"			
repeat_region	31737..31839	/rpt_family="Alu"			
repeat_region	31993..32138	/rpt_family="Alu"			
repeat_region	32744..32943	/rpt_family="MER1_type"			
repeat_region	33577..33663	/rpt_family="Alu"			
repeat_region	33610..33750	/rpt_family="B4"			
repeat_region	36229..36350	/rpt_family="Alu"			
repeat_region	36679..36762	/rpt_family="MIR"			
repeat_region	37830..37953	/rpt_family="MER1_type"			
repeat_region	38461..38565	/rpt_family="B2"			
repeat_region	38950..39327	/rpt_family="MaLR"			
repeat_region	39497..39639	/rpt_family="B4"			
repeat_region	40408..40684	/rpt_family="B4"			
repeat_region	41391..42306	/rpt_family="ERVK"			
repeat_region	43473..43832	/rpt_family="L1"			
repeat_region	44689..44835	/rpt_family="Tc2"			
repeat_region	45866..46200	/rpt_family="MaLR"			
repeat_region	46640..46786	/rpt_family="Alu"			
repeat_region	51588..51858	/rpt_family="MaLR"			
repeat_region	52268..52352	/rpt_family="ID"			
trNA	52276..52347	/products="tRNA-Ala"			
		/note="Likely pseudogene (HMM Sc=16.12 / Sec struct Sc=4.29)"			
repeat_region	54145..54483	/rpt_family="MaLR"			
repeat_region	54657..54860				
Query Match	2.0%;	Score 27; DB 10; Length 204855;			
Best Local Similarity	100.0%;	Pred. No. 0.00056;			
Matches 27; Conservative	0; Mismatches 0; Indels 0; Gaps 0;				
LOCUS	AC099108	232802 bp	DNA	linear	HTG 10-MAY-2003
DEFINITION	Rattus norvegicus clone CH230-186G8, WORKING DRAFT SEQUENCE.				
ACCESSION	AC099108	7	GI:30522525		
VERSION	HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.				
KEYWORDS	Rattus norvegicus (Norway rat)				
SOURCE	Rattus norvegicus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
REFERENCE	1 (bases 1 to 232802)				
AUTHORS	Muzny,D,Marle., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J., Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguitano,D., Ayalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H., Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F., Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M., Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E., Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Chen,Z., Chu,J., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Cress,A., D'Souza,L., Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Guevara,W., Davila,M.L., Davis,C., Derynck,C., Ding,Y., Dinh,H., Divya,K., Delgado,O., Denson,S., Deramo,C., Durbin,K., Duval,B., Eaves,K., Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Falls,T., Fan,G., Egan,A., Escotto,M., Eugene,C., Evans,C.A., Foster,M., Foster,P., Fernandez,S., Finley,M., Flagg,N., Forbes,L., Garner,T., Garza,M., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Grady,M., Guerra,W., Gebregorgis,E., Geer,K., Gill,R., Grady,M., Hamilton,C., Hamilton,K., Gunaratne,P., Haaland,W., Hamill,C., Henderson,N., Hernandez,J., Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idiebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorensuhewa,L., Loulsegid,H., Lozano,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Minet,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K., Nwackeleme,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfankoch,C., Plopper,F., Plopper,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rivas,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S., Shen,H., Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabot,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K., Valas,R., Vera,V., Villalana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlezyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.				
TITLE	Unpublished				
REFERENCE	2 (bases 1 to 232802)				
AUTHORS	Worley,K.C				
TITLE	Direct Submission				

JOURNAL

Submitted (09-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE

3 (bases 1 to 232802)

AUTHORS

Rat Genome Sequencing Consortium.

TITLE

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

JOURNAL

On May 10, 2003 this sequence version replaced gi:2309531.

COMMENT

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GGFQ

Center clone name: CH230-186G8

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 224396 bases at least Q40

Consensus quality: 225844 bases at least Q30

Consensus quality: 226943 bases at least Q20

Estimated insert size: 233584; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 232802: contig of 232802 bp in length.

FEATURES

source

1. 232802

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-186G8"

1. 1210

/note="wgs end extension

clone_end:Sp6"

1261..3361

/note="wgs end extension

clone_end:Sp6"

4020..6873

/note="wgs end extension

clone_end:Sp6"

8435..8986

/note="clone boundary

clone_end:Sp6

site:EcoRI

end_sequence:BH292506"

complement(230253..230668)

/note="clone boundary

clone_end:T7

site:EcoRI
end_sequence:BH292505"

ORIGIN

Query Match 2.0%; Score 27; DB 2; Length 232802;
Best Local Similarity 100.0%; Pred. No. 0.00057;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1301 AGATGTCACAAAAA 1327

DB 74169 AGATGTCACAAAAA 74195

Search completed: September 17, 2004, 04:46:08
Job time : 5170 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 16, 2004, 20:16:53 ; Search time 552 Seconds
(without alignments)
10212.607 Million cell updates/sec

Title: US-09-927-091-3_COPY_2500_3826
Perfect score: 1327
Sequence: 1 tccacagtgtcacagtagt.....aaaaaaaaaaaaaaaaaa 1327

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

- Database :
- 1: Geneseqn_29Jan04.*
 - 2: Geneseqn1980s.*
 - 3: Geneseqn2000s.*
 - 4: Geneseqn2001as.*
 - 5: Geneseqn2001bs.*
 - 6: Geneseqn2002s.*
 - 7: Geneseqn2003as.*
 - 8: Geneseqn2003bs.*
 - 9: Geneseqn2003cs.*
 - 10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1327	100.0	3826	6	ABK12806 Human cDN
2	1308	98.6	5858	4	AAK80624 Human imm
3	1308	98.6	22893	6	ABK12810 Human tum
4	1307	98.5	3243	9	ADC37268 Nuclear f
5	1257	94.7	5866	6	AAK80625 Human imm
6	1257	94.7	30676	6	ABK12811 Human tum
7	1256	94.6	3436	4	AAH14509 Human cDN
8	726	54.7	5858	4	AAK80626 Human imm
9	726	54.7	30625	6	ABK12808 Human tum
10	452	34.1	2207	9	ADC37454 Nuclear f
11	452	34.1	2207	9	ADC37264 Nuclear f
12	423	31.9	2246	9	ADC37266 Nuclear f
13	423	31.9	557	4	AAH11818 Human cDN
14	60	4.5	60	6	ABN43449 Human spl
15	29	2.2	614	6	ABN63808 Human can
16	27	2.0	470	3	AAK59529 Human sec
17	27	2.0	189013	7	ACF62741 Cancer ba
18	27	2.0	189013	7	ADB20856 MRP1 base
19	27	2.0	189013	9	ADB87945 Human UGT
20	27	2.0	189013	9	ADB87945 Human UGT
21	27	2.0	189013	9	ADB96928 Human MDR
22	26	2.0	273	6	ABQ60500 Human col
23	26	2.0	393	7	ABX41090 Bovine ES

97	25	1.9	4740	3	AAA35032	Aaa35032 Human ade
98	25	1.9	4740	3	AAF21154	Aaf21154 Human low
99	25	1.9	4740	6	ABK50046	Abk50046 cDNA enco
100	25	1.9	4740	6	ABK83766	Abk83766 Human cDN
101	25	1.9	4740	7	ABZ96848	Abz96848 Human nuc
102	25	1.9	4740	7	ABV76142	Abv76142 Human Mac
103	25	1.9	4740	9	ADD25614	Add25614 Binding d
104	25	1.9	5307	8	ADA38055	Ada38055 DNA relat
105	25	1.9	5363	8	ADA38051	Ada38051 DNA relat
106	25	1.9	5371	7	ABZ35979	Abz35979 Human sec
107	25	1.9	5684	9	ABZ72685	Abz72685 Human end
108	25	1.9	5808	7	ABX76355	Abx76355 Lung can
109	25	1.9	6405	5	AAF97850	Aaf97850 Human neu
110	25	1.9	7047	3	AAA35034	Aaa35034 Human ade
111	25	1.9	7047	7	AAF21156	Aaf21156 Human low
112	25	1.9	7047	7	ABZ96850	Abz96850 Human nuc
113	25	1.9	15765	6	AAD32327	Aad32327 Human NFK
114	25	1.9	15765	6	AAD32179	Aad32179 Human NFK
115	25	1.9	18099	8	AAD58283	Aad58283 Human tum
116	25	1.9	27884	4	AAK77781	Aak77781 Human imm
117	25	1.9	79528	6	AAI50814	Aai50814 Human can
118	25	1.9	110000	6	ABZ55320_0	Abz55320 Human tra
119	25	1.9	110000	3	ABZ55320_3	Abz55320 Human tra
120	25	1.9	226475	8	AAD58279	Aad58279 Human tum
121	25	1.9	341511	6	ABZ55200	Abz55200 Genomic D
122	24	1.8	40	4	AAH20345	Aah20345 HHV6 viru
123	24	1.8	42	4	AAH89926	Aah89926 Mycobacte
124	24	1.8	90	2	AAQ98109	Aaq98109 Comb body
125	24	1.8	92	6	ABV96450	Abv96450 Human pan
126	24	1.8	158	6	ABL86941	Ab186941 Human ova
127	24	1.8	166	4	AAH71409	Aah71409 Human cer
128	24	1.8	173	3	ABX46137	Abx46137 Bovine ES
129	24	1.8	203	5	ABV19399	Abv19399 Human pro
130	24	1.8	206	6	ABL69722	Ab169722 Prostate
131	24	1.8	206	6	ABN96234	Abn96234 Gene #273
132	24	1.8	229	7	ABX54617	Abx54617 Bovine ES
133	24	1.8	274	4	AAF59571	Aaf59571 Rat urote
134	24	1.8	282	5	ABV14508	Abv14508 Human pro
135	24	1.8	289	4	AAZ23972	Aaz23972 Human ova
136	24	1.8	289	5	AAH82537	Aah82537 Human ova
137	24	1.8	299	2	AAV96628	Aav96628 Human sec
138	24	1.8	312	5	ABA14476	Abal4476 Human ner
139	24	1.8	314	5	ABV06723	Abv06723 Human pro
140	24	1.8	333	5	ABV49172	Abv49172 Human pro
141	24	1.8	338	4	AAH09999	Aah09999 Human cDN
142	24	1.8	361	7	ABZ88882	Abz88882 Human nuc
143	24	1.8	368	6	ABT08933	Abt08933 Phase-1 R
144	24	1.8	372	6	ABL84265	Ab184265 Human ova
145	24	1.8	373	6	ABG85428	Abg85428 Arabidops
146	24	1.8	379	5	ABV18002	Abv18002 Human pro
147	24	1.8	386	4	AAI92258	Aai92258 Human pol
148	24	1.8	390	5	ABV54455	Abv54455 Human pro
149	24	1.8	391	5	ABV47791	Abv47791 Human pro
150	24	1.8	393	5	ABV36670	Abv36670 Human pro

ALIGNMENTS

RESULT 1
ABK12806
ID ABK12806 standard; cDNA; 3826 BP.

XX ABK12806;

XX AC

XX 18-JUN-2002 (first entry)

XX Human cDNA encoding tumour suppressor CAR-1.

XX

KW Human; ss; tumour suppressor; gene; CAR-1; cytostatic; cancer; tumour;

KW Gene therapy; brain cancer; lung cancer; liver cancer; kidney cancer;

KW Colon cancer; stomach cancer; breast cancer; endometrial cancer;

KW prostate cancer; testicular cancer; ovarian cancer; skin cancer;

KW head and neck cancer; oesophageal cancer; bone marrow cancer;
 KW chromosome 1p31-1p36.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT 604..2031
 FT CDS /*tag= a
 FT /product= "CAR-1"
 FT
 XX WO200212285-A2.
 XX 14-FEB-2002.
 XX
 XX 09-AUG-2001; 2001WO-US025269.
 XX
 XX 10-AUG-2000; 2000US-0225033P.
 XX 23-AUG-2000; 2000US-0227560P.
 XX
 XX (TEXA) UNIV TEXAS SYSTEM.
 PA
 XX Killary A, Chandler D, Lott S;
 PI WPI; 2002-269088/31.
 XX P-PSDB; AAU78657.
 DR
 XX New tumor suppressor CAR-1 polypeptides and polynucleotides, useful for
 PT diagnosing cancer, for altering the phenotype of a tumor cell, for
 PT treating cancers or as a diagnostic or prognostic indicator of cancer.
 PT
 XX Claim 2; Page 134-135; 185pp; English.
 CC
 CC The invention relates to an isolated polynucleotide encoding a
 CC polypeptide being tumour suppressor. CAR-1. Also included are fragments
 CC of the polynucleotide from 15-5000 nucleotides, fragments of the protein
 CC from 10-50 amino acids, an expression cassette comprising the
 CC polynucleotide under the control of a promoter operable in eukaryotic
 CC cells, a method for suppressing growth of a cancer cells by contacting
 CC the cells with the expression cassette (i.e. gene therapy), a cell
 CC comprising the expression cassette, an anti-CAR-1 monoclonal or
 CC polyclonal antibody, a hybridoma cell that produces the monoclonal
 CC antibody, a method of diagnosing a cancer by assessing the expression of
 CC CAR-1 tumour suppressor in the cells of a tissue sample from a subject,
 CC methods for altering the phenotype of a tumour cell, methods for treating
 CC a subject with cancer by administering the tumour suppressor CAR-1, or by
 CC administering a nucleic acid encoding the tumour suppressor CAR-1 and a
 CC promoter active in eukaryotic cells, where the promoter is operably
 CC linked to the region encoding the tumour suppressor, a non-human
 CC transgenic eukaryote lacking a functional CAR-1 gene, a non-human
 CC transgenic eukaryote that over-expresses CAR-1 as compared to a similar
 CC non-transgenic eukaryote, a method of screening a candidate substance for
 CC anti-tumour activity by contacting a cell lacking functional CAR-1
 CC polypeptide, with a candidate substance and determining the effect of the
 CC candidate substance on the cell, an anti-tumour composition produced by
 CC contacting a cell lacking functional CAR-1 polypeptide, with a candidate
 CC substance, determining the effect of the candidate substance on the cell,
 CC identifying a candidate inhibitor substance, and making a composition and
 CC an isolated and purified nucleic acid that hybridizes, under high
 CC stringency conditions, to a DNA segment comprising about 15-3826 bases of
 CC the CAR-1 cDNA. The polynucleotide is useful for diagnosing cancer, for
 CC altering the phenotype of a tumour cell, for treating cancers (e.g.,
 CC cancer of the brain, lung, liver, kidney, lymph node, pancreas, small
 CC intestine, blood cells, colon, stomach, breast, endometrium, prostate, or
 CC testicle, ovary, skin, head and neck, oesophagus, bone marrow, blood, or
 CC other tissues), and as a diagnostic or prognostic indicator of cancer.
 CC CAR-1 may also be used in screening compounds for activity in either
 CC stimulating CAR-1 activity, overcoming the lack of CAR-1 or blocking the
 CC effect of mutant CAR-1 molecule. The gene for CAR-1 is located on
 CC chromosome 1 (1p31-1p36). The present sequence is the cDNA encoding CAR-1
 XX
 XX SQ . Sequence 3826 BP; 797 A; 1286 C; 999 G; 744 T; 0 U; 0 Other;

Query Match

100.0%; Score 1327; DB 6; Length 3826;

Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1327; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	TCACAGTGTCTCAGGTAGTACTCTGCTCTAGGTTGCTGAGAGCAACCTCTCTGC	60
Db	2500	TCACAGTGTCTCAGGTAGTACTCTGCTCTAGGTTGCTGAGAGCAACCTCTCTGC	2559
QY	61	CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT	120
Db	2560	CACCCCCACACCAAGAACTATATGGTTCCTACTTCTCCCACTGATCTGCTGGTCAGTGAT	2619
QY	121	GATGCTGTGSCCTGTGGAAGGACCTGGTAGTGTAGTTCACACATTAATAGTCAATGSCCA	180
Db	2620	GATGCTGTGSCCTGTGGAAGGACCTGGTAGTGTAGTTCACACATTAATAGTCAATGSCCA	2679
QY	181	CCACCTTCTCTGCCACACAGGCGGAGACAGGTTAGGGTATACCCAAAGCTGATGACGAG	240
Db	2680	CCACCTTCTCTGCCACACAGGCGGAGACAGGTTAGGGTATACCCAAAGCTGATGACGAG	2739
QY	241	CCCATTAGCTTAAAGCAACTGACGAGCAAGCCCTCCCTGGATGATCGAGTCCCCAGTAG	300
Db	2740	CCCATTAGCTTAAAGCAACTGACGAGCAAGCCCTCCCTGGATGATCGAGTCCCCAGTAG	2739
QY	301	CTCTGAACAGAGTCCAGCCAAACCTCTTCAGCCAGGCTCTCTGACCTGCTAGGGTGCA	360
Db	2800	CTCTGAACAGAGTCCAGCCAAACCTCTTCAGCCAGGCTCTCTGACCTGCTAGGGTGCA	2859
QY	361	GGAGGCTTCCAGAAAGAGTGTGTATTAATAGGACCCACAGCACTGGAGGGGCTGTGGCT	420
Db	2860	GGAGGCTTCCAGAAAGAGTGTGTATTAATAGGACCCACAGCACTGGAGGGGCTGTGGCT	2919
QY	421	AGACCCCTTGTACAGCTTGGCACTATCTCAGTTAGGATCCTCTGCAGAAACCAAGAGC	480
Db	2920	AGACCCCTTGTACAGCTTGGCACTATCTCAGTTAGGATCCTCTGCAGAAACCAAGAGC	2979
QY	481	CACCTTGTAGTGTGTTAATAGACAGGATTTACTACCTGGCCCTCTGGTGGCTTGCAAAA	540
Db	2980	CACCTTGTAGTGTGTTAATAGACAGGATTTACTACCTGGCCCTCTGGTGGCTTGCAAAA	3039
QY	541	TTGTTGGAAGAGCTGGAGAGAGAGTCTGCTGAATTTCCAGAACTCCCGAGGCGCAGAT	600
Db	3040	TTGTTGGAAGAGCTGGAGAGAGAGTCTGCTGAATTTCCAGAACTCCCGAGGCGCAGAT	3099
QY	601	TCATCATGTCTGTGTGACAGAGAAAGCTGCCCCATCTGCAGAAAGCCACTATGCCAGA	660
Db	3100	TCATCATGTCTGTGTGACAGAGAAAGCTGCCCCATCTGCAGAAAGCCACTATGCCAGA	3159
QY	661	AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCGACGGTCCGTGCGAGCAATAGATGTC	720
Db	3160	AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCGACGGTCCGTGCGAGCAATAGATGTC	3219
QY	721	CTGAGGCTTGGCCCTCTCCCACTTCACTCAGTTCCCAATCTAAATTTTCAAGAGATT	780
Db	3220	CTGAGGCTTGGCCCTCTCCCACTTCACTCAGTTCCCAATCTAAATTTTCAAGAGATT	3279
QY	781	CTGTTTGGGGAACTTAAGTCAGATCCAGAACTTGGCTGCAAGGAGTCTGGGAAATCT	840
Db	3280	CTGTTTGGGGAACTTAAGTCAGATCCAGAACTTGGCTGCAAGGAGTCTGGGAAATCT	3339
QY	841	CAATTTCCCTAGAGGAAGTTAGGTTGGGTGGAGCAAGCCCACTGCTTTTCTGCCAC	900
Db	3340	CAATTTCCCTAGAGGAAGTTAGGTTGGGTGGAGCAAGCCCACTGCTTTTCTGCCAC	3399
QY	901	AGCATCCCAATCTGAGAGACTGGGAGAGGTGGATCCACATCTAGGTTGTCTGCCC	960
Db	3400	AGCATCCCAATCTGAGAGACTGGGAGAGGTGGATCCACATCTAGGTTGTCTGCCC	3459
QY	961	CTTGGCTCTATCCCTGCCAGAGGTGGGAACTGGAGAGTGGGCTGCAAGCTGAGCCTA	1020
Db	3460	CTTGGCTCTATCCCTGCCAGAGGTGGGAACTGGAGAGTGGGCTGCAAGCTGAGCCTA	3519
QY	1021	AATGCTCTCCCGGCTTGACTTTTCTTCTAGTCTCTGGGCTTAGATTCTGCACTTGGG	1080

PR	14-AUG-2000;	2000US-0225759P.	PR	08-NOV-2000;	2000US-0246611P.
PR	18-AUG-2000;	2000US-0226279P.	PR	08-NOV-2000;	2000US-0246613P.
PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0249207P.
PR	23-AUG-2000;	2000US-0226868P.	PR	17-NOV-2000;	2000US-0249208P.
PR	23-AUG-2000;	2000US-0227180P.	PR	17-NOV-2000;	2000US-0249209P.
PR	23-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249210P.
PR	30-AUG-2000;	2000US-0228242P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249217P.
PR	08-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249218P.
PR	08-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249445P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249464P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249300P.
PR	12-SEP-2000;	2000US-0232081P.	PR	01-DEC-2000;	2000US-0250160P.
PR	12-SEP-2000;	2000US-0231968P.	PR	01-DEC-2000;	2000US-0250391P.
PR	14-SEP-2000;	2000US-0232337P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232338P.	PR	05-DEC-2000;	2000US-0251989P.
PR	14-SEP-2000;	2000US-0232339P.	PR	06-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251869P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251989P.
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251990P.
PR	21-SEP-2000;	2000US-0234223P.	PR	11-DEC-2000;	2000US-0254037P.
PR	21-SEP-2000;	2000US-0234224P.	PR	05-JAN-2001;	2001US-0259678P.
PR	25-SEP-2000;	2000US-0234997P.	XX	(HUMA-) HUMAN GENOME SCI INC.	
PR	25-SEP-2000;	2000US-0234998P.	XX	Rosen CA, Barash SC, Ruben SM;	
PR	26-SEP-2000;	2000US-0235484P.	XX	WPI; 2001-483426/52.	
PR	27-SEP-2000;	2000US-0235834P.	XX	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PR	27-SEP-2000;	2000US-0235836P.	XX	useful for preventing, diagnosing and/or treating cancers and metastasis.	
PR	29-SEP-2000;	2000US-0236327P.	XX	Disclosure; SEQ ID NO 35436; 3071pp + Sequence Listing; English.	
PR	29-SEP-2000;	2000US-0236372P.	XX	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)	
PR	29-SEP-2000;	2000US-0236372P.	XX	amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic	
PR	29-SEP-2000;	2000US-0236368P.	XX	activity, and can be used in gene therapy and vaccine production. (I)	
PR	29-SEP-2000;	2000US-0236369P.	XX	proteins and polynucleotides may be used in the prevention, diagnosis and	
PR	29-SEP-2000;	2000US-0236370P.	XX	treatment of diseases associated with inappropriate (I) expression. For	
PR	29-SEP-2000;	2000US-0236370P.	XX	example, they may be used to treat disorders associated with decreased	
PR	02-OCT-2000;	2000US-0236802P.	XX	expression by rectifying mutations or deletions in a patient's genome	
PR	02-OCT-2000;	2000US-0237037P.	XX	that affect the activity of (I) by expressing inactive proteins or to	
PR	02-OCT-2000;	2000US-0237038P.	XX	supplement the patient's own production of (I). Additionally, (I)	
PR	02-OCT-2000;	2000US-0237039P.	XX	nucleic acids into a host cell and culturing the cell to prevent,	
PR	02-OCT-2000;	2000US-0237040P.	XX	protein. (I) proteins and polynucleotides may be used to prevent,	
PR	13-OCT-2000;	2000US-0239935P.	XX	diagnose and treat immune/haematopoietic-related diseases, especially	
PR	13-OCT-2000;	2000US-0239937P.	XX	cancers and cancer metastases of haematopoietic-derived cells. AAK64703	
PR	20-OCT-2000;	2000US-0240960P.	XX	to AAK87694 represent human immune/haematopoietic antigen genomic	
PR	20-OCT-2000;	2000US-0241221P.	XX	sequences from the present invention. AAK54942 to AAK54950 and AAK62169	
PR	20-OCT-2000;	2000US-0241785P.	XX	represent sequences used in the exemplification of the present invention	
PR	20-OCT-2000;	2000US-0241786P.	XX	Sequence 5858 BP; 1455 A; 1579 C; 1490 G; 1334 T; 0 U; 0 Other;	
PR	20-OCT-2000;	2000US-0241787P.	XX	Query Match 98.6%; Score 1308; DB 4; Length 5858;	
PR	20-OCT-2000;	2000US-0241809P.	XX	Best Local Similarity 100.0%; Pred.No. 0;	
PR	20-OCT-2000;	2000US-0241825P.	XX	Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
PR	01-NOV-2000;	2000US-0244617P.	XX	1 TCCACAGTCTCACAGGTAGTACCTGGTCTCCTAGGTGCTCCAGAGCCAACTCTCCCTGC 60	
PR	08-NOV-2000;	2000US-0246474P.	XX		
PR	08-NOV-2000;	2000US-0246475P.	XX		
PR	08-NOV-2000;	2000US-0246476P.	XX		
PR	08-NOV-2000;	2000US-0246477P.	XX		
PR	08-NOV-2000;	2000US-0246478P.	XX		
PR	08-NOV-2000;	2000US-0246523P.	XX		
PR	08-NOV-2000;	2000US-0246524P.	XX		
PR	08-NOV-2000;	2000US-0246525P.	XX		
PR	08-NOV-2000;	2000US-0246526P.	XX		
PR	08-NOV-2000;	2000US-0246527P.	XX		
PR	08-NOV-2000;	2000US-0246528P.	XX		
PR	08-NOV-2000;	2000US-0246532P.	XX		
PR	08-NOV-2000;	2000US-0246609P.	XX	1027 TCCACAGTCTCACAGGTAGTACCTGGTCTCCTAGGTGCTCCAGAGCCAACTCTCCCTGC 1086	
PR	08-NOV-2000;	2000US-0246610P.	XX		

QY 61 CACCCACACCAAGAACTATATGTTCTTCTACTTCTCCACTGATCTGCTGTCAGTGAT 120
Db 1087 CACCCACACCAAGAACTATATGTTCTTCTACTTCTCCACTGATCTGCTGTCAGTGAT 1146
QY 121 GATCTGTGGCCTGTGAAGCACTGTGTAGTGTGAGTCCACACATATAGTCAATGTCCTCA 180
Db 1147 GATCTGTGGCCTGTGAAGCACTGTGTAGTGTGAGTCCACACATATAGTCAATGTCCTCA 1266
QY 181 CCACCTTCTGCCACACAGGCGGAGGACAGGCTCAGGCTATACCAAGCTGATGTCAGAG 240
Db 1207 CCACCTTCTGCCACACAGGCGGAGGACAGGCTCAGGCTATACCAAGCTGATGTCAGAG 1266
QY 241 CCATATAGCTATAAGCAACTGCAGGACAGGCTCCTGTGATGATCGAGTCCCCAGTAG 300
Db 1267 CCATATAGCTATAAGCAACTGCAGGACAGGCTCCTGTGATGATCGAGTCCCCAGTAG 1326
QY 301 CTCTGACAGAGTCCAGCAACCTCTTTCAGCAGGCTCTGTGACCTGTAGGTTGCA 360
Db 1327 CTCTGACAGAGTCCAGCAACCTCTTTCAGCAGGCTCTGTGACCTGTAGGTTGCA 1386
QY 361 GGAGGCTTCCAGAGCAAGTGTGTATATAGGACCCCAAGCACTGGAGGGCTGTGGCT 420
Db 1387 GGAGGCTTCCAGAGCAAGTGTGTATATAGGACCCCAAGCACTGGAGGGCTGTGGCT 1446
QY 421 AGACCCCTGTGACACTTGCATCTATCTCAGTTAGGATCTGTGACAGAAACAGAGC 480
Db 1447 AGACCCCTGTGACACTTGCATCTATCTCAGTTAGGATCTGTGACAGAAACAGAGC 1506
QY 481 CACTGTAGTGTGTTAAATAGACAGGATTTACTCTGCCCCCTGGTGGCTTGCAAAA 540
Db 1507 CACTGTAGTGTGTTAAATAGACAGGATTTACTCTGCCCCCTGGTGGCTTGCAAAA 1566
QY 541 TTGTTGGAAGAGCTGGAGAGCAGACTCTGCTCAATTTCCAGAACTCCAGGCGCAGAT 600
Db 1567 TTGTTGGAAGAGCTGGAGAGCAGACTCTGCTCAATTTCCAGAACTCCAGGCGCAGAT 1626
QY 601 TCATCATGTCTGTGTGACAGGAAAGCTGCCGCCATCTGCAGGAAGCACTATGCGAGA 660
Db 1627 TCATCATGTCTGTGTGACAGGAAAGCTGCCGCCATCTGCAGGAAGCACTATGCGAGA 1686
QY 661 AGCTCTGACTGACAGAACTAGCTCCCTCTGTGCAACGCTGCGTGGCCAGCAATAGATGTC 720
Db 1687 AGCTCTGACTGACAGAACTAGCTCCCTCTGTGCAACGCTGCGTGGCCAGCAATAGATGTC 1746
QY 721 CTGAGGCTGCCCCCTCTCCCACTTCACTCAGTTCCCAATCTAAATTTTACAGAGAT 780
Db 1747 CTGAGGCTGCCCCCTCTCCCACTTCACTCAGTTCCCAATCTAAATTTTACAGAGAT 1806
QY 781 CTGTTGGGGAACTTAAGTCAAGTCCAGAACCTTGGCTGCAAGGAGTCTGGAAATCT 840
Db 1807 CTGTTGGGGAACTTAAGTCAAGTCCAGAACCTTGGCTGCAAGGAGTCTGGAAATCT 1866
QY 841 CATTTCCCTAGAGGAAGTTAGGTTGGGTGGAGCAAGCCCACTGCTGTTTTCTGCCCAC 900
Db 1867 CATTTCCCTAGAGGAAGTTAGGTTGGGTGGAGCAAGCCCACTGCTGTTTTCTGCCCAC 1926
QY 901 AGCATCCCAATCTGGAAGAACTCGGAGAGGGTGGAGTCCACATCTAGGTTGCTGCCCC 960
Db 1927 AGCATCCCAATCTGGAAGAACTCGGAGAGGGTGGAGTCCACATCTAGGTTGCTGCCCC 1986
QY 961 CTGTGCTCTATCCCTGCCAGAGGTGGGAACCTGGAGAGTGGGCTGCAAGACTGAGCCCTA 1020
Db 1987 CTGTGCTCTATCCCTGCCAGAGGTGGGAACCTGGAGAGTGGGCTGCAAGACTGAGCCCTA 2046
QY 1021 AATGCTCCCCGGCTTGACTTTCTTCTTCTAGTCTGGGCTTAGATTCTGCACTTGGGG 1080
Db 2047 AATGCTCCCCGGCTTGACTTTCTTCTTCTAGTCTGGGCTTAGATTCTGCACTTGGGG 2106
QY 1081 TCTCTGACACACACATCCCAAGTAGCCGAGAGCTTAAACACAGGGGGTCTTAA 1140
Db 2107 TCTCTGACACACACATCCCAAGTAGCCGAGAGCTTAAACACAGGGGGTCTTAA 2166

QY 1141 AATGCTGCCCGCCGACACCGGGCTCCTTGGCCAAAAGAAATGTCAGCCCTACCCCA 1200
Db 2167 AATGCTGCCCGCCGACACCGGGCTCCTTGGCCAAAAGAAATGTCAGCCCTACCCCA 2226
QY 1201 ACCCTTCAACTACAGAACTCTGGCCACCCCGAGTATTTTATTTAAATGTTGCCCA 1260
Db 2227 ACCCTTCAACTACAGAACTCTGGCCACCCCGAGTATTTTATTTAAATGTTGCCCA 2286
QY 1261 TTTTATGAGTATGATCAATTTGATTTATTAATTAAGTTACAGATGTCA 1308
Db 2287 TTTTATGAGTATGATCAATTTGATTTATTAATTAAGTTACAGATGTCA 2334
RESULT 3
ABK12810
ID ABK12810 standard; DNA; 22893 BP.
XX
AC ABK12810;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human tumour suppressor CAR-1, BAC clone RP11-150F21 3' sequence.
XX
DE Human; ds; tumour suppressor; CAR-1; cytostatic; cancer; tumour;
KW gene therapy; brain cancer; lung cancer; liver cancer; kidney cancer;
KW colon cancer; stomach cancer; breast cancer; endometrial cancer;
KW prostate cancer; testicular cancer; ovarian cancer; skin cancer;
KW head and neck cancer; oesophageal cancer; bone marrow cancer; BAC;
KW bacteria artificial chromosome; chromosome 1p31-1p36.
XX
OS Homo sapiens.
XX
XX
PH Key Location/Qualifiers
FT misc_feature 5820
FT /*tag= a
FT /note= "Nucleotides 5821-6360 of the present sequence as
FT nucleotide 5821 of the specification are illegible or missing,
FT nucleotide 6361 of the present sequence corresponds to
FT nucleotide 6361 of sequence as printed in the
FT specification"
XX
PN WO200212285-A2.
XX
XX 14-FEB-2002.
XX
XX 09-AUG-2001; 2001WO-US025269.
XX
XX 10-AUG-2000; 2000US-0225033P.
PR 23-AUG-2000; 2000US-0227560P.
XX
XX (TEXA) UNIV TEXAS SYSTEM.
XX
XX Killary A, Chandler D, Lott S;
PI WPI; 2002-269088/31.
DR
XX
PT New tumor suppressor CAR-1 polypeptides and polynucleotides, useful for
PT diagnosing cancer, for altering the phenotype of a tumor cell, for
PT treating cancers or as a diagnostic or prognostic indicator of cancer.
XX
PS Disclosure; Page 170-176; 185pp; English.
XX
XX The invention relates to an isolated polynucleotide encoding a
CC polypeptide being tumour suppressor, CAR-1. Also included are fragments
CC of the polynucleotide from 15-5000 nucleotides, fragments of the protein
CC from 10-50 amino acids, an expression cassette comprising the
CC polynucleotide under the control of a promoter operable in eukaryotic
CC cells, a method for suppressing growth of a cancer cells by contacting
CC the cells with the expression cassette (i.e. gene therapy), a cell
CC comprising the expression cassette, an anti-CAR-1 monoclonal or
CC polyclonal antibody, a hybridoma cell that produces the monoclonal
CC antibody, a method of diagnosing a cancer by assessing the expression of
CC CAR-1 tumour suppressor in the cells of a tissue sample from a subject,

CC methods for altering the phenotype of a tumour cell, methods for treating
CC a subject with cancer by administering the tumour suppressor CAR-1, or by
CC administering a nucleic acid encoding the tumour suppressor CAR-1 and a
CC promoter active in eukaryotic cells, where the promoter is operably
CC linked to the region encoding the tumour suppressor, a non-human
CC transgenic eukaryote lacking a functional CAR-1 gene, a non-human
CC transgenic eukaryote that over-expresses CAR-1 as compared to a similar
CC non-transgenic eukaryote, a method of screening a candidate substance for
CC anti-tumour activity by contacting a cell lacking functional CAR-1
CC polypeptide, with a candidate substance and determining the effect of the
CC candidate substance on the cell, an anti-tumour composition produced by
CC contacting a cell lacking functional CAR-1 polypeptide, with a candidate
CC substance, determining the effect of the candidate substance on the cell,
CC identifying a candidate inhibitor substance, and making a composition and
CC an isolated and purified nucleic acid that hybridizes, under high
CC stringency conditions, to a DNA segment comprising about 15-3826 bases of
CC the CAR-1 cDNA. The polynucleotide is useful for diagnosing cancer, for
CC altering the phenotype of a tumour cell, for treating cancers (e.g.
CC cancer of the brain, lung, liver, kidney, lymph node, pancreas, small
CC intestine, blood cells, colon, stomach, breast, endometrium, prostate,
CC testicle, ovary, skin, head and neck, oesophagus, bone marrow, blood, or
CC other tissues), and as a diagnostic or prognostic indicator of cancer.
CC CAR-1 may also be used in screening compounds for activity in either
CC stimulating CAR-1 activity, overcoming the lack of CAR-1 or blocking the
CC effect of mutant CAR-1 molecule. The gene for CAR-1 is located on
CC chromosome 1 (1p31-1p36). The present sequence is a BAC (bacterial
CC artificial chromosome) containing part of the CAR-1 gene
XX
SQ Sequence 22893 BP; 5720 A; 5989 C; 5664 G; 5415 T; 0 U; 105 Other;

Query Match 98.6%; Score 1308; DB 6; Length 22893;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCCACAGTGGTACAGGTTAGTACCTGGTCTAGGTTGCTTGGAGGCAACCTCTCTGC 60
Db 13010 TCCACAGTGGTACAGGTTAGTACCTGGTCTAGGTTGCTTGGAGGCAACCTCTCTGC 13069

Qy 61 CACCCCAACACAGAACTATAGTCTTCTCCACCTGATCTGCTGGTCAAGAT 120
Db 13070 CACCCCAACACAGAACTATAGTCTTCTCCACCTGATCTGCTGGTCAAGAT 13129

Qy 121 GATGCTGGTGGTGGAGGCACTGTAGTTCAGTCCACATTAAGTCAATGTGCA 180
Db 13130 GATGCTGGTGGTGGAGGCACTGTAGTTCAGTCCACATTAAGTCAATGTGCA 13189

Qy 181 CCACCTTCTGCCACAGGCCGAGGACAGGCTAGGTTATACCAAAAGCTGATSCAG 240
Db 13190 CCACCTTCTGCCACAGGCCGAGGACAGGCTAGGTTATACCAAAAGCTGATSCAG 13249

Qy 241 CCCATTAGCTTAAAGCACTCAGGCAAGCCTCCCTGGATGATCGAGTCCCAGTAG 300
Db 13250 CCCATTAGCTTAAAGCACTCAGGCAAGCCTCCCTGGATGATCGAGTCCCAGTAG 13309

Qy 301 CTCTGAAACAGAGTCCAGCCACCTCTTCAGCCAGGCTCTGTACCTGCTAGGTTGCA 360
Db 13310 CTCTGAAACAGAGTCCAGCCACCTCTTCAGCCAGGCTCTGTACCTGCTAGGTTGCA 13369

Qy 361 GGAGGCTTCCAGAGCAGTGTGTAAATTAGGACCCCAAGCACTGGGAGGGGCTGTGGCT 420
Db 13370 GGAGGCTTCCAGAGCAGTGTGTAAATTAGGACCCCAAGCACTGGGAGGGGCTGTGGCT 13429

Qy 421 AGACCCCTTGTGAGCTTGGATCTATCTCAGTATGATCTCTGCTGAGAAACAGAGC 480
Db 13430 AGACCCCTTGTGAGCTTGGATCTATCTCAGTATGATCTCTGCTGAGAAACAGAGC 13489

Qy 481 CACTTGTAGCTGGTTTAAATTAGACAGGATTTACTACCTGGCCCTGGTGGCTTGCAAA 540
Db 13490 CACTTGTAGCTGGTTTAAATTAGACAGGATTTACTACCTGGCCCTGGTGGCTTGCAAA 13549

Qy 541 TTGTTGGAAAGAGCTGAGAGAGCAGACTCTGCTGAATTTCCAGGAATCCCAAGCCAGAT 600
Db 13550 TTGTTGGAAAGAGCTGAGAGAGCAGACTCTGCTGAATTTCCAGGAATCCCAAGCCAGAT 13609

Qy 601 TCATCATGTCTGTGTGACACAGGAAAGCTGCCCCCATCTGCAGAAAGCCACTATGCCAGA 660
Db 13610 TCATCATGTCTGTGTGACACAGGAAAGCTGCCCCCATCTGCAGAAAGCCACTATGCCAGA 13669

Qy 661 AAGCTGCTGACTCAGCAAACTAGGCTCCCTCTGCTGCACGGTCCGTCGACCAAGCAATAGATGTC 720
Db 13670 AAGCTGCTGACTCAGCAAACTAGGCTCCCTCTGCTGCACGGTCCGTCGACCAAGCAATAGATGTC 13729

Qy 721 CTGAGGCTGCCCCCTCTCCCACTTCACTAGTTCACAAATCTAAATTTTACAGAGATT 780
Db 13730 CTGAGGCTGCCCCCTCTCCCACTTCACTAGTTCACAAATCTAAATTTTACAGAGATT 13789

Qy 781 CTGTTTGGGGAACTTAAAGTCAGATCCAGAAACCTTGGCTCAAGGGAGTCTGGGAAATGT 840
Db 13790 CTGTTTGGGGAACTTAAAGTCAGATCCAGAAACCTTGGCTCAAGGGAGTCTGGGAAATGT 13849

Qy 841 CATTTCCCTAGAAGGAACTTAGGTTAGGTTGGAGGAGCAAGCCCCACCTGGTTTCTGCGAC 900
Db 13850 CATTTCCCTAGAAGGAACTTAGGTTAGGTTGGAGGAGCAAGCCCCACCTGGTTTCTGCGAC 13909

Qy 901 AGCATCAATCGTCAGAACTCGGGAGAGGTGGAGTCCACATCTAGGGTTGTCTCTGCC 960
Db 13910 AGCATCAATCGTCAGAACTCGGGAGAGGTGGAGTCCACATCTAGGGTTGTCTCTGCC 13969

Qy 961 CTTGGCTCTATCCCTGCTCCAGAGGTGGAACTCGAGGAGTGGGCTGCAAGACTGAGCCTA 1020
Db 13970 CTTGGCTCTATCCCTGCTCCAGAGGTGGAACTCGAGGAGTGGGCTGCAAGACTGAGCCTA 14029

Qy 1021 AATGTCTCCCGGCTTGAATCTTTCTTAGTCTCTGGGCTAGATTCGCACTTGGG 1080
Db 14030 AATGTCTCCCGGCTTGAATCTTTCTTAGTCTCTGGGCTAGATTCGCACTTGGG 14089

Qy 1081 TCTGTGACACACACACATCCCAAGTGTAGCGGAGAGCTTAAACACAGGGGGTCTTTAA 1140
Db 14090 TCTGTGACACACACACATCCCAAGTGTAGCGGAGAGCTTAAACACAGGGGGTCTTTAA 14149

Qy 1141 AATGGCTGCCCGCCCAACCCGGGCTCCCTTGGGCAAAAGGAATGTGACGCTTACC 1200
Db 14150 AATGGCTGCCCGCCCAACCCGGGCTCCCTTGGGCAAAAGGAATGTGACGCTTACC 14209

Qy 1201 ACCCTTCAACTPACAGAACTTGGGCGCACCCAGCAGTATTTTATTTAAATGTGCCCA 1260
Db 14210 ACCCTTCAACTPACAGAACTTGGGCGCACCCAGCAGTATTTTATTTAAATGTGCCCA 14269

Qy 1261 TTTTATGAGTTATGATCAATTTGATTAATAATTAAGTTACAGATGCTCA 1308
Db 14270 TTTTATGAGTTATGATCAATTTGATTAATAATTAAGTTACAGATGCTCA 14317

RESULT 4
ADC37268
ID ADC37268 standard; DNA; 3243 BP.
XX
AC ADC37268;
XX
DT 18-DEC-2003 (first entry)
XX
DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 101.
XX
KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
KW cancer; infectious disease; bone disease; AIDS;
KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;
KW immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-Hiv;
KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
XX
OS Homo sapiens.
XX
EN WO2003048202-A2.
XX
PD 12-JUN-2003.
XX
PF 03-DEC-2002; 2002WO-JP012644.

XX 03-DEC-2001; 2001JP-00368692.
 PR 05-DEC-2001; 2001US-0335829P.
 PR 03-OCT-2002; 2002JP-00291302.
 PR 04-OCT-2002; 2002US-0415769P.
 XX
 PA (ASAH) ASAMI KASEI KK.
 XX
 PI Matsuda A, Muramatsu S;
 XX WPI: 2003-505282/47.
 DR P-PSDB; ADC37269.
 XX
 PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
 PT useful for treating inflammation, autoimmune diseases, cancers,
 PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
 PT ischemic disorders.
 XX
 PS Claim 4; SEQ ID NO 101; 938pp; English.
 XX
 CC The present invention relates to novel proteins and their coding
 CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
 CC kappaB). The proteins and their coding sequences are useful for treating
 CC a disease associated with NF-kappaB activation, such as inflammation,
 CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
 CC neurodegenerative diseases, or ischaemic disorders.
 XX
 SQ Sequence 3243 BP; 715 A; 1023 C; 863 G; 642 T; 0 U; 0 Other;

Query Match 98.5%; Score 1307; DB 9; Length 3243;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1307; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCACAGTGTGCAGGTAGTACCTGCTCCTAGGTTGCTGAGAGCCACCTCTCTGC 60
 DB 1937 TCACAGTGTGCAGGTAGTACCTGCTCCTAGGTTGCTGAGAGCCACCTCTCTGC 1996

QY 61 CACCCCCACACCAAGAACTATATGTTCTTCTTCTCCACTGATCTGCTGGTCACTGAT 120
 DB 1997 CACCCCCACACCAAGAACTATATGTTCTTCTTCTCCACTGATCTGCTGGTCACTGAT 2056

QY 121 GATGCTGTGGCTGTGGAGGACCTGCTGAGTGTGCTGAGAGCCACCTCTCTGC 180
 DB 2057 GATGCTGTGGCTGTGGAGGACCTGCTGAGTGTGCTGAGAGCCACCTCTCTGC 2116

QY 181 CCACCTTCTCTGCCACAGGCGGAGGACAGGCTGAGTGTGCTGAGAGCCACCTCTCTGC 240
 DB 2117 CCACCTTCTCTGCCACAGGCGGAGGACAGGCTGAGTGTGCTGAGAGCCACCTCTCTGC 2176

QY 241 CCACCTTCTCTGCCACAGGCGGAGGACAGGCTGAGTGTGCTGAGAGCCACCTCTCTGC 300
 DB 2177 CCACCTTCTCTGCCACAGGCGGAGGACAGGCTGAGTGTGCTGAGAGCCACCTCTCTGC 2236

QY 301 CTCTGACAGAGTCCAGGACACCTCTCTGAGGAGGCTCTGAGAGGCTCTGAGGAGGCT 360
 DB 2237 CTCTGACAGAGTCCAGGACACCTCTCTGAGGAGGCTCTGAGAGGCTCTGAGGAGGCT 2296

QY 361 GGAGGCTTCCAGAGAGTGTGTAATTTAGGACCAAGCACTGAGGAGGCTGTGGCT 420
 DB 2297 GGAGGCTTCCAGAGAGTGTGTAATTTAGGACCAAGCACTGAGGAGGCTGTGGCT 2356

QY 421 AGACCCCTTGTGAGACTTGGCACTATCTCAGTGTAGGATCTCTGCTGCAAAACCAAGAGC 480
 DB 2357 AGACCCCTTGTGAGACTTGGCACTATCTCAGTGTAGGATCTCTGCTGCAAAACCAAGAGC 2416

QY 481 CACTTGTAGCTGTTTAAATAGACAGAGATTTACTACCTGGCCCTCTGCTTGCAGAA 540
 DB 2417 CACTTGTAGCTGTTTAAATAGACAGAGATTTACTACCTGGCCCTCTGCTTGCAGAA 2476

QY 541 TTGTTGGAGAGCTGAGAGAGAGACTCTGCTGAAATTTCCAGAACTCCAGCGCCAGAT 600
 DB 2477 TTGTTGGAGAGCTGAGAGAGAGACTCTGCTGAAATTTCCAGAACTCCAGCGCCAGAT 2536

QY 601 TCATCATGTCTGTGTGACACAGGAAGCTGCCCCCATCTCTGACGAGGACCACTATGCCAGA 660
 DB 2537 TCATCATGTCTGTGTGACACAGGAAGCTGCCCCCATCTCTGACGAGGACCACTATGCCAGA 2596

QY 661 AAGCTGCTGACTGACAGAACTAGGCTCCCTCTGCGACGCTCGTGCAGCCCAATAGATGTC 720
 DB 2597 AAGCTGCTGACTGACAGAACTAGGCTCCCTCTGCGACGCTCGTGCAGCCCAATAGATGTC 2656

QY 721 CTGAGGCTGCCCCCTCTCCCACTTCTCACTCAGTTTCCCAATCTAAATTTTACAGAGATT 780
 DB 2657 CTGAGGCTGCCCCCTCTCCCACTTCTCACTCAGTTTCCCAATCTAAATTTTACAGAGATT 2716

QY 781 CTGTTTGGGGAATTAAGTCAGATCCAGAACTTGGCTGCAAGGGAGTCTGGAAATGT 840
 DB 2717 CTGTTTGGGGAATTAAGTCAGATCCAGAACTTGGCTGCAAGGGAGTCTGGAAATGT 2776

QY 841 CATTTCCTAGAGGAAGTTAGGTTGGTGGAGCAAGCCCACTGCTGTTTCTGCGAC 900
 DB 2777 CATTTCCTAGAGGAAGTTAGGTTGGTGGAGCAAGCCCACTGCTGTTTCTGCGAC 2836

QY 901 AGCATCAATCTGTAAGAACTCCGGAGAGGGTGGAGTCCACATCTAGGGTTGCTGCCCC 960
 DB 2837 AGCATCAATCTGTAAGAACTCCGGAGAGGGTGGAGTCCACATCTAGGGTTGCTGCCCC 2896

QY 961 CTTGGCTCTATCCCTGCCAGAGTGGGAACCTGAGGAGTGGCTGCAAGCTGAGCCTA 1020
 DB 2897 CTTGGCTCTATCCCTGCCAGAGTGGGAACCTGAGGAGTGGCTGCAAGCTGAGCCTA 2956

QY 1021 AATGTCTCTCCCGGCTTGAATTTCTTCTTAGTCTGGGGCTTAGATTCTGCACTTTGGG 1080
 DB 2957 AATGTCTCTCCCGGCTTGAATTTCTTCTTAGTCTGGGGCTTAGATTCTGCACTTTGGG 3016

QY 1081 TCTCTGACACACACACCACTCCCAAGTAGCCGGAAGAGCTAAACACAGGGGTTCTTAA 1140
 DB 3017 TCTCTGACACACACACCACTCCCAAGTAGCCGGAAGAGCTAAACACAGGGGTTCTTAA 3076

QY 1141 AATGGTGTCCCGGCTCCACCGGGCTCTCTTGGGCAAAAGGAATTTGACGCCCTACCCCA 1200
 DB 3077 AATGGTGTCCCGGCTCCACCGGGCTCTCTTGGGCAAAAGGAATTTGACGCCCTACCCCA 3136

QY 1201 ACCCTTCAACTACCAAGAACTCTGGGCCACCCCAAGCAGTATTTTATTAAGTTGTCGCCA 1260
 DB 3137 ACCCTTCAACTACCAAGAACTCTGGGCCACCCCAAGCAGTATTTTATTAAGTTGTCGCCA 3196

QY 1261 TTTTATGAGTTATGATCAATTTGTATTAATTAAGTTACAGATGTC 1307
 DB 3197 TTTTATGAGTTATGATCAATTTGTATTAATTAAGTTACAGATGTC 3243

RESULT 5
 ID AAK80625 standard; DNA; 5865 BP.
 AC AAK80625;
 XX
 XX
 DT 07-NOV-2001 (first entry)
 XX
 DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35437.
 XX
 XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
 KW cytostatic; gene therapy; vaccine; metastasis; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200157182-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001354.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 14-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218230P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 18-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229503P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0233397P.
 PR 14-SEP-2000; 2000US-0233398P.
 PR 14-SEP-2000; 2000US-0233399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0234848P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 06-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 XX WPI; 2001-483426/52.
 DR
 XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX
 XX Disclosure; SEQ ID NO 35437; 3071pp + Sequence Listing; English.
 PS
 XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
 CC

CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patient's own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention
 XX
 SQ Sequence 5866 BP; 1464 A; 1580 C; 1489 G; 1333 T; 0 U; 0 Other;

Query Match 94.7%; Score 1257; DB 4; Length 5866;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1307; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCACACAGTGTACAGTACTAGTCTAGGTTGCTGAGAGCAACCTCTCTGTC 60
 Db 1026 TCACACAGTGTACAGTACTAGTCTAGGTTGCTGAGAGCAACCTCTCTGTC 1085

QY 61 CACCCCCACACCAAGAACTATATGGTTCCTACTTCCCACTGATCTGCTGTCAGTGAT 120
 Db 1086 CACCCCCACACCAAGAACTATATGGTTCCTACTTCCCACTGATCTGCTGTCAGTGAT 1145

QY 121 GATGCTGTGGCTGTGGAGGCACTGTAGTGTAGTTCACACATATAGTCAATGTGCCA 180
 Db 1146 GATGCTGTGGCTGTGGAGGCACTGTAGTGTAGTTCACACATATAGTCAATGTGCCA 1205

QY 181 GCACCTTCTGCCCCACAGGCGGAGGACAGGGTACGGTATATCCCAAGCTGATGACAG 240
 Db 1206 GCACCTTCTGCCCCACAGGCGGAGGACAGGGTACGGTATATCCCAAGCTGATGACAG 1265

QY 241 CCATTAGCTTAAGCACTGTCAGACCAAGCCCTCCCTGATGATCGAGTCCCAAGTAG 300
 Db 1266 CCATTAGCTTAAGCACTGTCAGACCAAGCCCTCCCTGATGATCGAGTCCCAAGTAG 1325

QY 301 CTCTGAACAGAGTCCAGCAACCCCTCTTCCAGCAGGCTCTGTGACCTGTAGGCTGCA 360
 Db 1326 CTCTGAACAGAGTCCAGCAACCCCTCTTCCAGCAGGCTCTGTGACCTGTAGGCTGCA 1385

QY 361 GGAGGCTTCCAGACAGTGTGTATATAGGACCAAGCACTGGAGGGGCTGTGGCT 420
 Db 1386 GGAGGCTTCCAGACAGTGTGTATATAGGACCAAGCACTGGAGGGGCTGTGGCT 1445

QY 421 AGACCCCTTGTGACACTTGGCATCTATCTCAGTTAGGATCCCTGCTCAGAAACAGAGC 480
 Db 1446 AGACCCCTTGTGACACTTGGCATCTATCTCAGTTAGGATCCCTGCTCAGAAACAGAGC 1505

QY 481 CACTGTAGCTGGTTTAATATAGACAGGATTTACTACCTGGCCCTGTGGCTTGCAAA 540
 Db 1506 CACTGTAGCTGGTTTAATATAGACAGGATTTACTACCTGGCCCTGTGGCTTGCAAA 1565

QY 541 TTGTTGGAGAGCTGGAGAGCAGACTCTGCTGAATTTCCAGAACTCCAGCGCCAGAT 600
 Db 1566 TTGTTGGAGAGCTGGAGAGCAGACTCTGCTGAATTTCCAGAACTCCAGCGCCAGAT 1625

QY 601 TCATCATGTCTGTGTGACAGAAAGCTGCCCATCTGTCAGGAGCCACTATGCCAGA 660
 Db 1626 TCATCATGTCTGTGTGACAGAAAGCTGCCCATCTGTCAGGAGCCACTATGCCAGA 1685

QY 661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGTCAGCGGTCCGTGCCAGCAATAGATGTC 720
 Db 1686 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGTCAGCGGTCCGTGCCAGCAATAGATGTC 1745

QY 721 CTGAGGCTTGGCCCTCTCCCACTTCACTCAGTTCCCAATCTAAATTTTCAAGAGATT 780

Db 1746 CTGAGGCTGCCCTCTCCCACTTCCACTCAGTCCCAAACTAAATTTTACAGAGATT 1805
 QY 781 CTGTTTGGGGAACCTTAAGTCAAGTCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 840
 Db 1806 CTGTTTGGGGAACCTTAAGTCAAGTCCAGAACCTTGGCTGCAAGGGAGTCTGGGAAATGT 1865
 QY 841 CATTTCCCTAGAGGAGTGTAGGGTGGTGGAGCAAGCCCACTGCTGGTTTCTGCGAC 900
 Db 1866 CATTTCCCTAGAGGAGTGTAGGGTGGTGGAGCAAGCCCACTGCTGGTTTCTGCGAC 1925
 QY 901 AGCATCCAAATCGTGAAGAACTCCGGAGAGGGTGGAGTCCACATCTAGGGTTGCTCGGCC 960
 Db 1926 AGCATCCAAATCGTGAAGAACTCCGGAGAGGGTGGAGTCCACATCTAGGGTTGCTCGGCC 1985
 QY 961 CTTGGCTCTATCCCTGCCCAGAGGTTGGGAACTGGAGAGTGGCTGCAAGACTGAGCCTA 1020
 Db 1986 CTTGGCTCTATCCCTGCCCAGAGGTTGGGAACTGGAGAGTGGCTGCAAGACTGAGCCTA 2045
 QY 1021 AATGTCTCCCGGCTTGAATTTTCTTCTAGTCCCTGGGCTAGATCTGCACTTGGGG 1080
 Db 2046 AATGTCTCCCGGCTTGAATTTTCTTCTAGTCCCTGGGCTAGATCTGCACTTGGGG 2105
 QY 1081 TCTCTGACACAAACACACCATCCCAAGTAGCCGGAAGAGCTAAACACAGGGGTTCTTAA 1140
 Db 2106 TCTCTGACACAAACACACCATCCCAAGTAGCCGGAAGAGCTAAACACAGGGGTTCTTAA 2165
 QY 1141 AATGGCTGCCCGCCACACCGGGCTCCCTTGGGCAAAAGAAATGTGACCCCTAGCCCA 1200
 Db 2166 AATGGCTGCCCGCCACACCGGGCTCCCTTGGGCAAAAGAAATGTGACCCCTAGCCCA 2225
 QY 1201 ACCCTTCAACTACCAAGTCTGGCCACCCAGCAGTATTTTATTAAGTTACAGATGCA 1308
 Db 2226 ACCCTTCAACTACCAAGTCTGGCCACCCAGCAGTATTTTATTAAGTTACAGATGCA 2333
 QY 1261 TTTATGAGTTATGATCAATTTGTATTAATTAAGTTACAGATGCA 1308
 Db 2286 TTTATGAGTTATGATCAATTTGTATTAATTAAGTTACAGATGCA 2333

RESULT 6

ABK12811
 ID ABK12811 standard; DNA; 30676 BP.

XX AC ABK12811;

XX DT 18-JUN-2002 (first entry)

XX DE Human tumour suppressor CAR-1, BAC clone RP11-131M11.

XX KW Human; ds; tumour suppressor; CAR-1; cytostatic; cancer; tumour;
 KW gene therapy; brain cancer; lung cancer; liver cancer; kidney cancer;
 KW colon cancer; stomach cancer; breast cancer; endometrial cancer;
 KW prostate cancer; testicular cancer; ovarian cancer; skin cancer;
 KW head and neck cancer; oesophageal cancer; bone marrow cancer; BAC;
 KW bacteriella artificial chromosome; chromosome lp31-1p36.

XX OS Homo sapiens.

XX PN WO200212285-A2.

XX PD 14-FEB-2002.

XX PF 09-AUG-2001; 2001WO-US025269.

XX PR 10-AUG-2000; 2000US-0225033P.

XX PR 23-AUG-2000; 2000US-0227560P.

XX PA (TEXA) UNIV TEXAS SYSTEM.

XX PI Killary A, Chandler D, Lott S;

XX DR WPI; 2002-269088/31.

PT New tumor suppressor CAR-1 polypeptides and polynucleotides, useful for
PT diagnosing cancer, for altering the phenotype of a tumor cell, for
PT treating cancers or as a diagnostic or prognostic indicator of cancer.
XX Disclosure; Page 176-185; 185pp; English.
XX
XX The invention relates to an isolated polynucleotide encoding a
CC polypeptide being tumor suppressor, CAR-1. Also included are fragments
CC of the polynucleotide from 15-5000 nucleotides, fragments of the protein
CC from 10-50 amino acids, an expression cassette comprising the
CC polynucleotide under the control of a promoter operable in eukaryotic
CC cells, a method for suppressing growth of a cancer cells by contacting
CC the cells with the expression cassette (i.e. gene therapy), a cell
CC comprising the expression cassette, an anti-CAR-1 monoclonal or
CC polyclonal antibody, a hybridoma cell that produces the monoclonal
CC antibody, a method of diagnosing a cancer by assessing the expression of
CC CAR-1 tumor suppressor in the cells of a tissue sample from a subject,
CC methods for altering the phenotype of a tumor cell, methods for treating
CC a subject with cancer by administering the tumor suppressor CAR-1, or by
CC administering a nucleic acid encoding the tumor suppressor CAR-1 and a
CC promoter active in eukaryotic cells, where the promoter is operably
CC linked to the region encoding the functional CAR-1 gene, a non-human
CC transgenic eukaryote that over-expresses CAR-1 as compared to a similar
CC non-transgenic eukaryote, a method of screening a candidate substance for
CC anti-tumor activity by contacting a cell lacking functional CAR-1
CC polypeptide, with a candidate substance and determining the effect of the
CC candidate substance on the cell, an anti-tumor composition produced by
CC contacting a cell lacking functional CAR-1 polypeptide, with a candidate
CC substance, determining the effect of the candidate substance on the cell,
CC identifying a candidate inhibitor substance, and making a composition and
CC an isolated and purified nucleic acid that hybridizes, under high
CC stringency conditions, to a DNA segment comprising about 15-3826 bases of
CC the CAR-1 cDNA. The polynucleotide is useful for diagnosing cancer, for
CC altering the phenotype of a tumor cell, for treating cancers (e.g.,
CC cancer of the brain, lung, liver, kidney, lymph node, pancreas, small
CC intestine, blood cells, colon, stomach, breast, endometrium, prostate,
CC testicle, ovary, skin, head and neck, esophagus, bone marrow blood, or
CC other tissues), and as a diagnostic or prognostic indicator of cancer.
CC CAR-1 may also be used in screening compounds for activity in either
CC stimulating CAR-1 activity, overcoming the lack of CAR-1 or blocking the
CC effect of mutant CAR-1 molecule. The gene for CAR-1 is located on
CC chromosome 1 (1p31-1p36). The present sequence is a BAC (bacterial
CC artificial chromosome) containing part of the CAR-1 gene
XX
SQ Sequence 30676 BP; 7591 A; 7621 C; 7888 G; 6944 T; 0 U; 632 Other;

Query Match 94.7%; Score 1257; DB 6; Length 30676;
Best Local Similarity 99.9%; Pred No. 0;
Matches 1307; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TCCACAGGGTGCACAGGTAGTACCTGGTCTAGGGTGGCTGAGAGCCAACTCTCCCTGC 60
DB 24440 TCCACAGGGTGCACAGGTAGTACCTGGTCTAGGGTGGCTGAGAGCCAACTCTCCCTGC 24499

QY 61 CACCCCCACCAAGAACTATATGTTCTACTCTCCCACTGATCTGCTGCTAGTAT 120
DB 24500 CACCCCCACCAAGAACTATATGTTCTACTCTCCCACTGATCTGCTGCTAGTAT 24559

QY 121 GATGCTGGCTGGCTGGAGGACCTGGTGTAGTGGTCCACATATATAGTCAATGTGCCA 180
DB 24560 GATGCTGGCTGGCTGGAGGACCTGGTGTAGTGGTCCACATATATAGTCAATGTGCCA 24619

QY 181 CCACCTTCTGCTGCCACAGCCGAGGACAGGGTGGGTATACCCAAAGCTGATCCAGAG 240
DB 24620 CCACCTTCTGCTGCCACAGCCGAGGACAGGGTGGGTATACCCAAAGCTGATCCAGAG 24679

QY 241 CCCATTAGCTAAAGCACTGACAGCAAGCTCTCCCTGGATGATCGAGGCTCCCCAGTAG 300
DB 24680 CCCATTAGCTAAAGCACTGACAGCAAGCTCTCCCTGGATGATCGAGGCTCCCCAGTAG 24739

QY 301 CTCTGAACAAGAGTCCAGCAACCCCTCTTTCAGCAGGCTCTGTGACCTGTAGGGTGCA 360
DB 24740 CTCTGAACAAGAGTCCAGCAACCCCTCTTTCAGCAGGCTCTGTGACCTGTAGGGTGCA 24799

RESULT 7
AAH14509
ID AAH14509 standard; cDNA; 3436 BP.
XX
AC AAH14509;
XX

DB 24740 CTCTGAACAAGAGTCCAGCAACCCCTCTTTCAGCAGGCTCTGTGACCTGTAGGGTGCA 24799
QY 361 GGAGGCTTCCAGAGCAGTTGTTGTAATTTAGGACCCCAAGCACTGGGAGGGCTGTGGCT 420
DB 24800 GGAGGCTTCCAGAGCAGTTGTTGTAATTTAGGACCCCAAGCACTGGGAGGGCTGTGGCT 24859
QY 421 AGACCCCTTGTCCAGACTTGGCATCTATCTCAGTTAGGATCCTCTCTGCAGAAAAACAAGC 480
DB 24860 GGACCCCTTGTCCAGACTTGGCATCTATCTCAGTTAGGATCCTCTCTGCAGAAAAACAAGC 24919
QY 481 CACTTGTAGCTGTTTAATTTAGACAGAGTTACTACCTGGCCCTCTGCTGGCTTGCAGAAA 540
DB 24920 CACTTGTAGCTGTTTAATTTAGACAGAGTTACTACCTGGCCCTCTGCTGGCTTGCAGAAA 24979
QY 541 TTGTTGGAAGAGCTGGAGAGCAGACTCTCTGAAATTTCCAGGAATCTCCAGCCGCGAGAT 600
DB 24980 TTGTTGGAAGAGCTGGAGAGCAGACTCTCTGAAATTTCCAGGAATCTCCAGCCGCGAGAT 25039
QY 601 TCATCATGCTGTTGTTGACCAAGAAAGCTGCCCCCAATCTCCAGGAAGCCCAATATGCCAGA 660
DB 25040 TCATCATGCTGTTGTTGACCAAGAAAGCTGCCCCCAATCTCCAGGAAGCCCAATATGCCAGA 25099
QY 661 AAGCTGTGATCTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCAATAGATGTC 720
DB 25100 AAGCTGTGATCTGCAGAACTAGGCTCCCTCTGCCACGGTCCGTGCCAGCAATAGATGTC 25159
QY 721 CTGAGGCTGCCCTCTCCCACTTCACTCAGTTCCCAAAATCTAAATTTTACAGAGATT 780
DB 25160 CTGAGGCTGCCCTCTCCCACTTCACTCAGTTCCCAAAATCTAAATTTTACAGAGATT 25219
QY 781 CTGTTTGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCTCAGGAGTCTGGGAATGT 840
DB 25220 CTGTTTGGGGAACTTAAGTCAGATCCAGAACCTTGGCTGCTCAGGAGTCTGGGAATGT 25279
QY 841 CATTTCCCTAGAGGAAGTTAGGCTGGTGGAGCAAGCCCACTGGTTTCTTTCGCCAC 900
DB 25280 CATTTCCCTAGAGGAAGTTAGGCTGGTGGAGCAAGCCCACTGGTTTCTTTCGCCAC 25339
QY 901 AGCATCAATCGTAGAAGTCTCGGAGAGGGTGGAGTCCACATCTAGGGTTGTCTGCCC 960
DB 25340 AGCATCAATCGTAGAAGTCTCGGAGAGGGTGGAGTCCACATCTAGGGTTGTCTGCCC 25399
QY 961 CTTGGCTCTATPCCCTGCCAGAGTGGGAACCTGGAGAGTGGGTGCAAGACTGAGCCCTA 1020
DB 25400 CTTGGCTCTATPCCCTGCCAGAGTGGGAACCTGGAGAGTGGGTGCAAGACTGAGCCCTA 25459
QY 1021 AATGTCTCCCGGCTTGAATTTTCTTCTAGTCTCGGGCTAGATTCTGCACCTGGGG 1080
DB 25460 AATGTCTCCCGGCTTGAATTTTCTTCTAGTCTCGGGCTAGATTCTGCACCTGGGG 25519
QY 1081 TCTCTGACACAAACACACATCCCAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
DB 25520 TCTCTGACACAAACACACATCCCAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 25579
QY 1141 AATGGCTGCCCGCCACCCCGGGCTCCCTTTGGGCAAGAAATGTGAGCCCTTACCCCA 1200
DB 25580 AATGGCTGCCCGCCACCCCGGGCTCCCTTTGGGCAAGAAATGTGAGCCCTTACCCCA 25639
QY 1201 ACCCTTCAACTACAGATCTGGGCCACCCAGCAGTATTTTATTTAAATGTGGCCCA 1260
DB 25640 ACCCTTCAACTACAGATCTGGGCCACCCAGCAGTATTTTATTTAAATGTGGCCCA 25699
QY 1261 TTTTATGATTTATGATCAATTTGTTAAATTTAAAGTTTACAGATGCA 1308
DB 25700 TTTTATGATTTATGATCAATTTGTTAAATTTAAAGTTTACAGATGCA 25747

PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 01-DEC-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 06-DEC-2000; 2000US-0256719P.
 PR 08-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Rosen CA, Barash SC, Ruben SM;
 PI WPI; 2001-483426/52.
 XX
 DR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and metastasis.
 XX
 PS Disclosure; SEQ ID NO 35438; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention
 CC
 XX Sequence 5858 BP; 1481 A; 1575 C; 1462 G; 1340 T; 0 U; 0 Other;
 SQ
 Query Match 54.7%; Score 726; DB 4; Length 5858;
 Best Local Similarity 99.4%; Pred. No. 1.9e-267;
 Matches 1146; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
 QY 157 TCACACATTATAGTCATGTCGCCACCACTTCTGCCCCAGCCGAGGACAGGGTGAG 216
 DB 1182 TCACACATTATAGTCATGTCGCCACCACTTCTGCCCCAGCCGAGGACAGGGTGAG 1241
 QY 217 GGTATACCCAAAGCTGATGCAGAGCCATTAGCTAAAGCACTGCAGGCAAGCCCTCTTCAGCCAG 336
 DB 1242 GGTATACCCAAAGCTGATGCAGAGCCATTAGCTAAAGCACTGCAGGCAAGCCCTCTTCAGCCAG 336
 QY 277 CTGGATGATCGAGGTCCTCCAGTAGCTCTGAAACAAGAGTCGAGCAACCTCTTCAGCCAG 336
 DB 1302 CTGGATGATCGAGGTCCTCCAGTAGCTCTGAAACAAGAGTCGAGCAACCTCTTCAGCCAG 1361
 QY 337 GCCTCTGTACCTGTAGGTCGAGAGGCTCCGAGGAGCTGTGTGTAATTAGGACCC 396
 DB 1362 GCCTCTGTACCTGTAGGTCGAGAGGCTCCGAGGAGCTGTGTGTAATTAGGACCC 1421
 QY 397 AAGCACT-GGGAGGGGCTGTGGCTAGACCCCTGTGCAGACTGGCATCTATCTAGTTA 455
 DB 1422 AAGCACTGGGAGGGGCTGTGGCTAGACCCCTGTGTGCTAGACTTGGCATCTATCTAGTTA 1481

QY 456 GGATCCTCTGCGAGAAAACAGAGCCACTTGTAGCTGTTTAAATTTAGACAAGGATTTACT 515
 DB 1482 GGATCCTCTGCGAGAAAACAGAGCCACTTGTAGCTGTTTAAATTTAGACAAGGATTTACT 1541
 QY 516 ACCTGGCCCTGCTGGCTTGCAGAAATTTGTCGAGAGCTGAGAGAGAGCTCTGCTGAA 575
 DB 1542 ACCTGGCCCTGCTGGCTTGCAGAAATTTGTCGAGAGCTGAGAGAGAGCTCTGCTGAA 1601
 QY 576 TTTCAGGAACCTCCAGCGCCAGATTTCATATGCTGTGTGTGACCCAGGAAAGCTCCGCC 635
 DB 1602 TTTCAGGAACCTCCAGCGCCAGATTTCATATGCTGTGTGTGACCCAGGAAAGCTCCGCC 1661
 QY 636 ATCTGCGAGAGCCACTATGCCAGAAAGCTGCTGACTGCGAGAACTAGGCTCCCTCTGCCA 695
 DB 1662 ATCTGCGAGAGCCACTATGCCAGAAAGCTGCTGACTGCGAGAACTAGGCTCCCTCTGCCA 1721
 QY 696 CGGTCCCTGCGCAGCAATAGATGTCCTGAGGCCCTCCCTCTCCCTCTCCCTCTCAGTTCC 755
 DB 1722 CGGTCCCTGCGCAGCAATAGATGTCCTGAGGCCCTCCCTCTCCCTCTCAGTTCC 1781
 QY 756 CAAATCTAAATTTTACAAGAGATTCTGTTTGGGGAACTTAAGTCAGATCCAGAACTT 815
 DB 1782 CAAATCTAAATTTTACAAGAGATTCTGTTTGGGGAACTTAAGTCAGATCCAGAACTT 1841
 QY 816 GGCTGCGAGAGCTCTGGGAAATGTCATTTCCCTAGAGGAAAGTTAGGCTGGGTGGAGCA 875
 DB 1842 GGCTGCGAGAGCTCTGGGAAATGTCATTTCCCTAGAGGAAAGTTAGGCTGGGTGGAGCA 1901
 QY 876 AGCCCCACCTCGGCTTTTTCGCCACAGCATCCCAATCTGAGAACTCGGAGAGGGTGA 935
 DB 1902 AGCCCCACCTCGGCTTTTTCGCCACAGCATCCCAATCTGAGAACTCGGAGAGGGTGA 1961
 QY 936 GTCCACATCTAGGGTTGTCTGCCCCCTTGGCTCTATCCCTGCCAGAGGTGGGAACCTGA 995
 DB 1962 GTCCACATCTAGGGTTGTCTGCCCCCTTGGCTCTATCCCTGCCAGAGGTGGGAACCTGA 2021
 QY 996 GAGTGGGCTGCAAGACTGAGCCCTAAATGTCCTCCCGGCTTGAATTTCTTTCTAGTCC 1055
 DB 2022 GAGTGGGCTGCAAGACTGAGCCCTAAATGTCCTCCCGGCTTGAATTTCTTTCTAGTCC 2081
 QY 1056 TGGGGCTCTAGATCTGACCTTGGGCTCTGACACACACACCATCCCAAGTAGCCGGA 1115
 DB 2082 TGGGGCTCTAAATCTGACCTTGGGCTCTGACACACACACCATCCCAAGTAGCCGGA 2141
 QY 1116 AGAGCTAAACACAGGGGGTCTTAAATGGCTGCCCGCCGACCCGGGCTCCCTTTGGGC 1175
 DB 2142 AAAGCTAAACACAGGGGGTCTTAAATGGCTGCCCGCCGACCCGGGCTCCCTTTGGGC 2201
 QY 1176 AAAAGGAATGTGACGCCCTACCCCAACCTCTCAACTACAGAAATCTGGGCCACCCGAGCA 1235
 DB 2202 AAAAGGAATGTGACGCCCTACCCCAACCTCTCAACTACAGAAATCTGGGCCACCCGAGCA 2261
 QY 1236 GTATTTTATTTAAATTTGGCCATTTTATGAGTTATGATCAATTTGTTATTAATTA 1295
 DB 2262 GTATTTTATTTAAATTTGGCCATTTTATGAGTTATGATCAATTTGTTATTAATTA 2321
 QY 1296 GTTACAGATGCA 1308
 DB 2322 GTTACAGATGCA 2334

RESULT 9
 ABK12808
 ID ABK12808 standard; DNA; 30625 Bp.
 XX
 AC ABK12808;
 XX
 DT 18-JUN-2002 (first entry)
 XX Human tumour suppressor CAR-1, BAC clone 392H05 3' sequence.
 DE Human; ds; tumour suppressor; CAR-1; cytostatic; cancer; tumour;
 KW gene therapy; brain cancer; liver cancer; lung cancer; kidney cancer;
 KW

Db 24211 GTATTTTAAATGTTCCCATTTTATGAGTATGATCAATTTGTTATTAATAAA 24270
 Qy 1296 GTTACAGATGTCA 1308
 Db 24271 GTTACAGATGTCA 24283
 RESULT 10
 ADC37454
 ID ADC37454 standard; DNA; 2207 BP.
 AC ADC37454;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 287.
 XX
 KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
 KW cancer; infectious disease; bone disease; AIDS;
 KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;
 KW immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-Hiv;
 KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO2003048202-A2.
 XX
 PD 12-JUN-2003.
 XX
 PF 03-DEC-2002; 2002WO-JP012644.
 XX
 PR 03-DEC-2001; 2001JP-00368692.
 PR 05-DEC-2001; 2001US-0335829P.
 PR 03-OCT-2002; 2002JP-00291302.
 PR 04-OCT-2002; 2002US-0415769P.
 XX
 PA (ASAH) ASahi KASEI KK.
 XX
 PI Matsuda A, Muramatsu S;
 XX
 DR WPI; 2003-505282/47.
 XX
 DR P-PSDB; ADC37455.
 XX
 PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
 PT useful for treating inflammation, autoimmune diseases, cancers,
 PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
 PT ischemic disorders.
 XX
 PS Claim 4; SEQ ID NO 287; 938pp; English.
 XX
 CC The present invention relates to novel proteins and their coding
 CC sequences (ADC37168-ADC37453), which activate nuclear factor kappa B (NF-
 kappaB). The proteins and their coding sequences are useful for treating
 CC a disease associated with NF-kappaB activation, such as inflammation,
 CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
 CC neurodegenerative diseases, or ischaemic disorders.
 XX
 SQ Sequence 2207 BP; 414 A; 784 C; 618 G; 391 T; 0 U; 0 Other;
 Query Match 34.1%; Score 452; DB 9; Length 2207;
 Best Local Similarity 99.8%; Pred. No. 7.4e-163;
 Matches 502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 801 CAGATCCAGACCTTGGCTGCAAGGAGTCTGGAAATGTCTATCCCTAGAGGAAGTT 860
 Db 1705 CAGATCCAGACCTTGGCTGCAAGGAGTCTGGAAATGTCTATCCCTAGAGGAAGTT 1764
 Qy 861 AGGGTGGGTGGAGCAAGCCACCTCGTTTCTGCCACAGATCCATCGTAGAGAC 920
 Db 1765 AGGGTGGGTGGAGCAAGCCACCTCGTTTCTGCCACAGATCCATCGTAGAGAC 1824
 Qy 921 TCGGGAGAGGGTGGAGTCCACATCTAGGGTGTGCTGCGCCCTTATCCCTGCCCA 980

Db 1825 TCGGAGAGGGTGGAGTCCATCTAGGGTGTCTGCCCTGGCTCTATCCCTGCCCA 1884
 Qy 981 GAGTGGGAACTGGAGAGTGGGCTGCAAGACTGAGCTAAATGTCTCCCGGCTTGAC 1040
 Db 1885 GAGTGGGAACTGGAGAGTGGGCTGCAAGACTGAGCTAAATGTCTCCCGGCTTGAC 1944
 Qy 1041 TTTTCTTTCTAGTCTGGGCTTAGATTCTGCACTTGGGCTCTCTGACACACACCAT 1100
 Db 1945 TTTTCTTTCTAGTCTGGGCTTAGATTCTGCACTTGGGCTCTCTGACACACACCAT 2004
 Qy 1101 CCAGAGTACCGGAGAGCTAAACACAGGGGTCTTAAATGGTCTGCCCGCCACCC 1160
 Db 2005 CCAGAGTACCGGAGAGCTAAACACAGGGGTCTTAAATGGTCTGCCCGCCACCC 2064
 Qy 1161 GGGCTCCCTTGGGCAAAAGAAATTGTGAGCCCTTACCCCAACCTTCACTACAGAAATC 1220
 Db 2065 GGGCTCCCTTGGGCAAAAGAAATTGTGAGCCCTTACCCCAACCTTCACTACAGAAATC 2124
 Qy 1221 TGGGCAACCCAGAGTATTTTATTTAAATGTGCCCCATTTATGAGTTATGATCAAT 1280
 Db 2125 TGGGCCACCCAGAGTATTTTATTTAAATGTGCCCCATTTATGAGTTATGATCAAT 2184
 Qy 1281 TTGTATTAAATTAAGTTACAGA 1303
 Db 2185 TTGTATTAAATTAAGTTACAGA 2207
 RESULT 11
 ADC37264
 ID ADC37264 standard; DNA; 2207 BP.
 XX
 AC ADC37264;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 97.
 XX
 KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
 KW cancer; infectious disease; bone disease; AIDS;
 KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;
 KW immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-Hiv;
 KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO2003048202-A2.
 XX
 PD 12-JUN-2003.
 XX
 PF 03-DEC-2002; 2002WO-JP012644.
 XX
 PR 03-DEC-2001; 2001JP-00368692.
 PR 05-DEC-2001; 2001US-0335829P.
 PR 03-OCT-2002; 2002JP-00291302.
 PR 04-OCT-2002; 2002US-0415769P.
 XX
 PA (ASAH) ASahi KASEI KK.
 XX
 PI Matsuda A, Muramatsu S;
 XX
 DR WPI; 2003-505282/47.
 XX
 DR P-PSDB; ADC37265.
 XX
 PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
 PT useful for treating inflammation, autoimmune diseases, cancers,
 PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
 PT ischemic disorders.
 XX
 PS Claim 4; SEQ ID NO 97; 938pp; English.
 XX
 CC The present invention relates to novel proteins and their coding
 CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-

CC kappaB). The proteins and their coding sequences are useful for treating
 CC a disease associated with NF-kappaB activation, such as inflammation,
 CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
 CC neurodegenerative diseases, or ischaemic disorders.
 XX
 SQ Sequence 2207 BP; 414 A; 784 C; 618 G; 391 T; 0 U; 0 Other;
 Query Match 34.1%; Score 452; DB 9; Length 2207;
 Best Local Similarity 99.8%; Pred. No. 7.4e-163;
 Matches 502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 801 CAGATCCAGAACCTTGGCTGCAAGGAGTCTGGAAATGTCATTTCCTAGAGGAAGTT 860
 Db 1705 CAGATCCAGAACCTTGGCTGCAAGGAGTCTGGAAATGTCATTTCCTAGAGGAAGTT 1764
 QY 861 AGGTGGGTGGAGCAAGCCCCACCTGGCTTTTTCGCCACAGCATCAATCGTGAAGAAC 920
 Db 1765 AGGTGGGTGGAGCAAGCCCCACCTGGCTTTTTCGCCACAGCATCAATCGTGAAGAAC 1824
 QY 921 TCGGGAGAGGTGGAGTCCACATCTAGGTTCCTGCGCCCTTGGCTCTATCCCTGCCCA 980
 Db 1825 TCGGGAGAGGTGGAGTCCACATCTAGGTTCCTGCGCCCTTGGCTCTATCCCTGCCCA 1884
 QY 981 GAGGTGGAACTGGAGGAGTGGGTGCAAGACTGAGCTAAATGTCCTCCCGGCTTGAC 1040
 Db 1885 GAGGTGGAACTGGAGGAGTGGGTGCAAGACTGAGCTAAATGTCCTCCCGGCTTGAC 1944
 QY 1041 TTTTCTTTCTAGTCTGGGCTAGATCTGCATCTGGGGTCTCTGACACAAACACCAT 1100
 Db 1945 TTTTCTTTCTAGTCTGGGCTAGATCTGCATCTGGGGTCTCTGACACAAACACCAT 2004
 QY 1101 CCCAAAGTAGCGGAGAGCTAAACACAGGGGGTCTTAAATGGCTGCCCGCCGCC 1160
 Db 2005 CCCAAAGTAGCGGAGAGCTAAACACAGGGGGTCTTAAATGGCTGCCCGCCGCC 2064
 QY 1161 GGGCTCCCTTGGGCAAAAGAAATGTGTCAGCCCTACCCCAACCCCTTCAACTACCAGATC 1220
 Db 2065 GGGCTCCCTTGGGCAAAAGAAATGTGTCAGCCCTACCCCAACCCCTTCAACTACCAGATC 2124
 QY 1221 TGGGCCACCCGAGCATATTTTATTTAAATGTTGGCCATTTATGAGTTATGATCAAT 1280
 Db 2125 TGGGCCACCCGAGCATATTTTATTTAAATGTTGGCCATTTATGAGTTATGATCAAT 2184
 QY 1281 TTGTATTAATTAAGTTTACAGA 1303
 Db 2185 TTGTATTAATTAAGTTTACAGA 2207
 RESULT 12
 ID ADC37266 standard; DNA; 2246 BP.
 XX
 AC ADC37266;
 XX
 DT 18-DEC-2003 (first entry)
 DE Nuclear factor kappa B (NF-kappaB) activating gene, SEQ ID 99.
 XX
 KW Nuclear factor kappa B; NF-kappaB; inflammation; autoimmune disease;
 KW cancer; infectious disease; bone disease; AIDS;
 KW neurodegenerative disease; ischaemic disorder; Antiinflammatory;
 KW Immunomodulator; Cytostatic; Antimicrobial; Osteopathic; Anti-HIV;
 KW Neuroprotective; Nootropic; Cardiant; Gene therapy; human; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO2003048202-A2.
 XX
 PD 12-JUN-2003.
 XX
 PF 03-DEC-2002; 2002WO-JP012644.
 XX
 PF 03-DEC-2001; 2001JP-00368692.
 PR

PR 05-DEC-2001; 2001US-0335829P.
 PR 03-OCT-2002; 2002JP-00291302.
 PR 04-OCT-2002; 2002US-0415769P.
 XX
 PA (ASAH) ASAH KASEI KK.
 XX
 XX Matsuda A, Muramatsu S;
 XX
 XX MPI; 2003-505282/47.
 DR P-PSDB; ADC37267.
 DR
 XX
 PT New purified protein that activates nuclear factor kappa B (NF-kappaB),
 PT useful for treating inflammation, autoimmune diseases, cancers,
 PT infectious diseases, bone diseases, AIDS, neurodegenerative diseases or
 PT ischaemic disorders.
 XX
 XX Claim 4; SEQ ID NO 99; 938pp; English.
 PS
 XX
 CC The present invention relates to novel proteins and their coding
 CC sequences (ADC37168-ADC37455), which activate nuclear factor kappa B (NF-
 CC kappaB). The proteins and their coding sequences are useful for treating
 CC a disease associated with NF-kappaB activation, such as inflammation,
 CC autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS,
 CC neurodegenerative diseases, or ischaemic disorders.
 XX
 SQ Sequence 2246 BP; 418 A; 805 C; 623 G; 400 T; 0 U; 0 Other;
 Query Match 34.1%; Score 452; DB 9; Length 2246;
 Best Local Similarity 99.8%; Pred. No. 7.4e-163;
 Matches 502; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 801 CAGATCCAGAACCTTGGCTGCAAGGAGTCTGGAAATGTCATTTCCTAGAGGAAGTT 860
 Db 1744 CAGATCCAGAACCTTGGCTGCAAGGAGTCTGGAAATGTCATTTCCTAGAGGAAGTT 1803
 QY 861 AGGTGGGTGGAGCAAGCCCCACCTGGCTTTTTCGCCACAGCATCAATCGTGAAGAAC 920
 Db 1804 AGGTGGGTGGAGCAAGCCCCACCTGGCTTTTTCGCCACAGCATCAATCGTGAAGAAC 1863
 QY 921 TCGGGAGAGGTGGAGTCCACATCTAGGTTCCTGCGCCCTTGGCTCTATCCCTGCCCA 980
 Db 1864 TCGGGAGAGGTGGAGTCCACATCTAGGTTCCTGCGCCCTTGGCTCTATCCCTGCCCA 1923
 QY 981 GAGGTGGGAACTGGAGGAGTGGGTGCAAGACTGAGCTAAATGTCCTCCCGGCTTGAC 1040
 Db 1924 GAGGTGGGAACTGGAGGAGTGGGTGCAAGACTGAGCTAAATGTCCTCCCGGCTTGAC 1983
 QY 1041 TTTTCTTTCTAGTCTGGGGCTAGATCTGCATCTGGGGTCTCTGACACACACACCAT 1100
 Db 1984 TTTTCTTTCTAGTCTGGGGCTAGATCTGCATCTGGGGTCTCTGACACACACACCAT 2043
 QY 1101 CCCAAAGTAGCGGAGAGCTAAACACAGGGGGTCTTAAATGGCTGCCCGCCGCC 1160
 Db 2044 CCCAAAGTAGCGGAGAGCTAAACACAGGGGGTCTTAAATGGCTGCCCGCCGCC 2103
 QY 1161 GGGCTCCCTTGGGCAAAAGAAATGTGTCAGCCCTACCCCAACCCCTTCAACTACCAGATC 1220
 Db 2104 GGGCTCCCTTGGGCAAAAGAAATGTGTCAGCCCTACCCCAACCCCTTCAACTACCAGATC 2163
 QY 1221 TGGGCCACCCGAGCATATTTTATTTAAATGTTGGCCATTTATGAGTTATGATCAAT 1280
 Db 2164 TGGGCCACCCGAGCATATTTTATTTAAATGTTGGCCATTTATGAGTTATGATCAAT 2223
 QY 1281 TTGTATTAATTAAGTTTACAGA 1303
 Db 2224 TTGTATTAATTAAGTTTACAGA 2246
 RESULT 13
 ID AAH11818/c
 XX AAH11818 standard; cDNA; 557 BP.
 AC AAH11818;

```

XX 26-JUN-2001 (first entry)
XX Human cDNA clone (3'-primer) SEQ ID NO:8653.
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EP1074617-A2.
XX 07-FEB-2001.
XX 28-JUL-2000; 2000EP-00116126.
XX 29-JUL-1999; 99JP-00248036.
XX 27-AUG-1999; 99JP-00300253.
XX 11-JAN-2000; 2000JP-00118776.
XX 02-MAY-2000; 2000JP-00183767.
XX 09-JUN-2000; 2000JP-00241899.
XX (HELI-) HELIX RES INST.
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX Primer sets for synthesising polynucleotides, particularly the 5602 full-
XX length cDNAs defined in the specification, and for the detection and/or
XX diagnosis of the abnormality of the proteins encoded by the full-length
XX cDNAs.
XX Claim 3; SEQ ID NO 8653; 2537pp + Sequence Listing; English.
XX The present invention describes primer sets for synthesising 5602 full-
XX length cDNAs defined in the specification. Where a primer set comprises:
XX (a) an oligo-dT primer and an oligonucleotide complementary to the
XX complementary strand of a polynucleotide which comprises one of the 5602
XX nucleotide sequences defined in the specification, where the
XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in the
XX specification. The primer sets can be used in antisense therapy and in
XX gene therapy. The primers are useful for synthesising polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
XX represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX oligonucleotides, all of which are used in the exemplification of the
XX present invention
XX Sequence 557 BP; 142 A; 133 C; 139 G; 136 T; 0 U; 7 Other;
XX Query Match 31.9%; Score 423; DB 4; Length 557;
XX Best Local Similarity 100.0%; Pred. No. 1.1e-151;
XX Matches 423; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX QY 885 TCGGTTTTTCTGCCACAGCATCCAAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATC 944
XX DB 423 TCGGTTTTTCTGCCACAGCATCCAAATCGTGAAGAACTCGGGAGAGGGTGGAGTCCACATC 364
XX QY 945 TAGGGTTGCTGCCCTTGGCTCTATCCCTGCCAGAGTGGAACTGGAGAGTGGGC 1004
XX DB 363 TAGGGTTGCTGCCCTTGGCTCTATCCCTGCCAGAGTGGAACTGGAGAGTGGGC 304
XX QY 1005 TGCAGAGACTGAGCCTAAATGTCTCCCGCGCTTGACTTTTCTTCTAGTCTGGGGCCCTA 1064

```

```

DB 303 TGCAGAGACTGAGCCTAAATGTCTCCCGCGCTTGACTTTTCTTCTAGTCTGGGGCTA 244
QY 1065 GATTCTGCACTTGGGGTCTCTGTGACACAAACACACCATCCCAAAGTAGCCGGAAGAGCTAAA 1124
DB 243 GATTCTGCACTTGGGGTCTCTGTGACACAAACACACCATCCCAAAGTAGCCGGAAGAGCTAAA 184
QY 1125 CACAGGGGGTCTTAAATGGCTGCCCGCCACCCCGGGCTCCCTTTGGGCAAAAGGAAT 1184
DB 183 CACAGGGGGTCTTAAATGGCTGCCCGCCACCCCGGGCTCCCTTTGGGCAAAAGGAAT 124
QY 1185 TGTGAGCCCTACCCCAACCCCTTCAACTACAGATCTGGGCCACCCACGAGTATTTTAA 1244
DB 123 TGTGAGCCCTACCCCAACCCCTTCAACTACAGATCTGGGCCACCCACGAGTATTTTAA 64
QY 1245 TTTAAATGTGCCCCATTTTATGAGTTATGATCAATTTGATTAAATTAAGTTACAGAT 1304
DB 63 TTTAAATGTGCCCCATTTTATGAGTTATGATCAATTTGATTAAATTAAGTTACAGAT 4
QY 1305 GTC 1307
DB 3 GTC 1

```

RESULT 14
ABN43449
ID ABN43449 standard; DNA; 60 BP.
XX AC ABN43449;
XX AC AC
DT 15-JUL-2002 (first entry)
XX DE Human spliced transcript detection oligonucleotide SEQ ID NO:16197.
XX KW Human; mouse; rat; splice transcript; detection; RNA transcript;
XX KW splice variant; transcriptome; oligonucleotide library; ss.
XX OS Homo sapiens.
XX PN WO200210449-A2.
XX PD 07-FEB-2002.
XX PF 20-JUL-2001; 2001WO-IB001903.
XX PR 28-JUL-2000; 2000US-0221607P.
XX PR 02-MAY-2001; 2001US-0287724P.
XX PA (COMP-) COMPUGEN INC.
XX PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX WPI; 2002-257383/30.
XX New oligonucleotide libraries comprising oligonucleotides which
XX selectively hybridize to mRNAs transcribed from a transcription unit of a
XX genome, useful for detecting tissue-, pathology-, and developmental-
XX specific genes.
XX Example 1; SEQ ID NO 16197; 47pp; English.
XX The present invention describes oligonucleotide libraries for detecting
XX messenger RNAs that populate a (sub-)transcriptome, where the (sub-
XX)transcriptome comprises messenger RNAs transcribed from multiple
XX transcription units that populate a genome. The library comprises several
XX oligonucleotides, each capable of hybridising selectively to a set of
XX messenger RNAs transcribed from a given transcription unit of the genome,
XX which encodes one or more messenger RNA splice variants. The
XX oligonucleotide libraries are useful for detecting mRNAs from a
XX biological sample, in expression profiling studies, in qualitatively or
XX quantitatively characterising the corresponding transcriptome, and in
XX detecting RNA transcripts and splice variants of human or animal
XX transcriptomes. The libraries may also be used as specialised mini

CC libraries to detect transcripts of a sub-transcriptome under a particular
CC biological or pathological state, and so allowing the detection of tissue
CC - and pathological-specific genes such as those genes only expressed in
CC specific tissue under a specific pathological condition; to detect
CC developmental specific genes; and to detect RNA transcripts and splice
CC variants of a transcriptome of a patient suffering from a particular
CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
CC rats, humans and mice, which are used in the exemplification of the
CC present invention. N.B. The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 60 BP; 17 A; 21 C; 11 G; 11 T; 0 U; 0 Other;

Query Match 4.5%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 4.6e-13;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1169 CTTGGCAAAAGGAATTGTACGCCCTACCCCAACCTTCACTACCAAGAAATCTGGGCAC 1228
Db 1 CTTGGCAAAAGGAATTGTACGCCCTACCCCAACCTTCACTACCAAGAAATCTGGGCAC 60

RESULT 15
ABN63808
ID ABN63808 standard; cDNA; 614 BP.

XX
AC ABN63808;

XX
DT 28-JUN-2002 (first entry)

XX
DE Human cancer related polynucleotide SEQ ID NO 3775.

XX
KW Human; cytostatic; gene expression; gene mapping; tissue profiling;
KW gene therapy; cancer; tumour; gene; ss.

XX
OS Homo sapiens.

XX
FN WO200214500-A2.

XX
PD 21-FEB-2002.

XX
PF 16-AUG-2001; 2001WO-US025840.

XX
PR 16-AUG-2000; 2000US-0226326P.

XX
PA (CHIR) CHIRON CORP.

XX
PA (HYSE-) HYSEQ INC.

XX
PI Escobedo J, Garcia PD, Sudduth-Klinger J, Reinhard C, Randazzo F;
PI Lamson G, Scott EM, Zhang G, Kassam A, Pot D, Labat I;

XX
DR WPI; 2002-241905/29.

XX
PT New nucleic acid for producing a polypeptide, detecting differentially
PT expressed genes correlated with a cancerous state of a mammalian cell,
PT and inhibiting tumor growth.

XX
PS Claim 1; SEQ ID NO 3775; 883pp + Sequence Listing; English.

XX
CC The invention relates to an isolated polynucleotide (ABN27253-ABN33262)
CC with cytostatic activity. The polynucleotide is used to produce a
CC polypeptide, to detect differentially expressed genes correlated with a
CC cancerous state of a mammalian cell and to inhibit tumor growth. The
CC polynucleotide is used as a probe in mapping and tissue profiling. The
CC encoded polypeptide and antibodies to the polypeptide can also be used
CC for therapeutic and diagnostic purposes. The polynucleotide is useful for
CC gene therapy. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 614 BP; 187 A; 127 C; 137 G; 163 T; 0 U; 0 Other;

Query Match 2.2%; Score 29; DB 6; Length 614;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 ACAGATGTCMAAAAAAAAAAAAAAAAAAAAAA 1327
Db 585 ACAGATGTCMAAAAAAAAAAAAAAAAAAAAAA 613

RESULT 16

AAC59529
ID AAC59529 standard; cDNA; 470 BP.

XX
AC AAC59529;

XX
DT 15-FEB-2001 (first entry)

XX
DE Human secreted protein cDNA sequence #23.

XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; ss.

XX
OS Homo sapiens.

XX
FN WO200055352-A2.

XX
PD 21-SEP-2000.

XX
PF 09-MAR-2000; 2000WO-US006044.

XX
PR 12-MAR-1999; 99US-0124099P.

XX
PR 03-DEC-1999; 99US-0168661P.

XX
PA (HUMA-) HUMAN GENOME SCI INC.

XX
PI Rosen CA, Ruben SM, Komatsoulis G;

XX
DR WPI; 2000-602124/57.

XX
DR P-ESDB; AAB34238.

XX
PT Novel human secreted proteins useful for diagnosis, prevention and
PT treatment of disorders including neurological, cell proliferative,
PT cardiovascular, autoimmune and inflammatory disorders and microbial
PT infections.

XX
PS Claim 1; Page 331; 383pp; English.

XX
CC The invention relates to the isolation of genes AAC59507-C59556 encoding
CC 50 human secreted proteins AAB34218-B34264. The genes can be used to
CC generate fusion proteins by linking to the gene for the human
CC immunoglobulin G Fc portion (AAC59498), for increasing the stability of
CC the fusion protein as compared to the human protein only. The genes and
CC proteins are useful for preventing, ameliorating or treating medical
CC conditions, e.g. by protein or gene therapy. The genes are isolated from
CC a range of human tissues disclosed in the specification. The nucleic
CC acids, proteins, antibodies and (ant)agonists are useful in the
CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b)
CC immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
XX infections

SQ Sequence 470 BP; 165 A; 90 C; 66 G; 149 T; 0 U; 0 Other;

Query Match 2.0%; Score 27; DB 3; Length 470;


```

Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1301 AGATGTCACAAAAAAGAAAAAAGAAAAA 1327
Db 410 AGATGTCACAAAAAAGAAAAAAGAAAAA 436

RESULT 17
ACF62741
ID ACF62741 standard; DNA; 189013 BP.
XX AC ACF62741;
XX DT 08-OCT-2003 (first entry)
XX DE Cancer based on CYP3A5 related polynucleotide SEQ ID NO:669.
XX KW Cancer; CYP3A5; irinotecan; pharmaceutical; malignant glioma;
XX KW cytochrome p450; subfamily IIIA; nifedipine oxidase; polypeptide 5;
XX KW cytosstatic; gene; ds.
XX OS Unidentified.
XX PN WO2003013534-A2.
XX PD 20-FEB-2003.
XX PF 23-JUL-2002; 2002WO-EP008219.
XX PR 23-JUL-2001; 2001EP-00117608.
XX PR 24-MAY-2002; 2002EP-00011710.
XX PA (EPID-) EPIDAUROS BIOLOGIE AG.
XX PI Heinrich G, Kerb R;
XX PS WPI; 2003-268144/26.
XX PT New use of irinotecan for preparation of compositions for treating cancer
XX PT in subject having genome with variant allele comprising cytochrome p450,
XX PT subfamily IIIA, polypeptide 5 polynucleotide, termed CYP3A5.
XX PS Disclosure; SEQ ID NO 669; 86pp; English.
XX CC The present invention describes the use of irinotecan (I) or its
XX CC derivative for the preparation of a pharmaceutical composition for
XX CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic
XX CC cancer, or malignant glioma in a subject having a genome with a variant
XX CC allele which comprises a cytochrome p450, subfamily IIIA (nifedipine
XX CC oxidase), polypeptide 5 (CYP3A5) polynucleotide (II). (I) and (II) have
XX CC cytostatic activity. The therapeutic applications of (I) is improved,
XX CC since it is possible to individually treat a subject with an appropriate
XX CC dosage and/or an appropriate derivative of (I). Therefore, undesirable,
XX CC harmful or toxic effects are efficiently avoided. Unnecessary and
XX CC potentially harmful treatment of those subjects who do not respond to the
XX CC treatment with substances (nonresponders), as well as the development of
XX CC drug resistances due to suboptimal drug dosing can be avoided. ACF62200
XX CC to ACF62751 and ABW34912 to ABW35013 represent sequences used in the
XX CC exemplification of the present invention
SQ Sequence 189013 BP; 53603 A; 38975 C; 39348 G; 57087 T; 0 U; 0 Other;

Query Match 2.0%; Score 27; DB 7; Length 189013;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1301 AGATGTCACAAAAAAGAAAAAAGAAAAA 1327
Db 123937 AGATGTCACAAAAAAGAAAAAAGAAAAA 123963

RESULT 18
ACF62741
ID ACF62741 standard; DNA; 189013 BP.
XX AC ACF62741;
XX DT 04-DEC-2003 (first entry)
XX DE Human UGT1A1 gene sequence SEQ ID NO:669.
XX KW irinotecan; cancer; UGT1A1; cytosstatic; topoisomerase I inhibitor;
XX KW colorectal cancer; cervical cancer; gastric cancer; lung cancer;

```

```

ADB20856
ID ADB20856 standard; DNA; 189013 BP.
XX AC ADB20856;
XX DT 20-NOV-2003 (first entry)
XX DE MRP1 based cancer related nucleic acid SEQ ID NO:669.
XX KW irinotecan; colorectal cancer; cervical cancer; gastric cancer;
XX KW lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;
XX KW variant allele; multidrug resistance protein 1; MRP1; cytosstatic; gene;
XX KW ds.
XX OS Unidentified.
XX PN WO2003013533-A2.
XX PD 20-FEB-2003.
XX PF 23-JUL-2002; 2002WO-EP008200.
XX PR 23-JUL-2001; 2001EP-00117608.
XX PR 24-MAY-2002; 2002EP-00011710.
XX PA (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.
XX PI Heinrich G, Kerb R;
XX PS WPI; 2003-354397/33.
XX PT Use of irinotecan or its derivative for preparation of a pharmaceutical
XX PT composition for treating cancer in a subject having a genome with a
XX PT variant allele comprising a multidrug resistance protein 1
XX PT polynucleotide.
XX PS Disclosure; SEQ ID NO 669; 100pp; English.
XX CC The present invention describes a method for the use of irinotecan (I) or
XX CC its derivative for the preparation of a pharmaceutical composition for
XX CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic
XX CC cancer, or malignant glioma in a subject having a genome with a variant
XX CC allele which comprises a multidrug resistance protein 1 (MRP1)
XX CC polynucleotide (II). (I) has cytostatic activity. (I) or its derivative
XX CC can be used for the preparation of a pharmaceutical composition for
XX CC treating colorectal, cervical, gastric, lung, ovarian or pancreatic
XX CC cancer, or malignant glioma in a subject, where the subject is a human
XX CC (preferably African or Asian) or a mouse. The present sequence represents
XX CC a sequence which is used in the exemplification of the present invention.
SQ Sequence 189013 BP; 53603 A; 38975 C; 39348 G; 57087 T; 0 U; 0 Other;

Query Match 2.0%; Score 27; DB 7; Length 189013;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1301 AGATGTCACAAAAAAGAAAAAAGAAAAA 1327
Db 123937 AGATGTCACAAAAAAGAAAAAAGAAAAA 123963

RESULT 19
ADB87945
ID ADB87945 standard; DNA; 189013 BP.
XX AC ADB87945;
XX DT 04-DEC-2003 (first entry)
XX DE Human UGT1A1 gene sequence SEQ ID NO:669.
XX KW irinotecan; cancer; UGT1A1; cytosstatic; topoisomerase I inhibitor;
XX KW colorectal cancer; cervical cancer; gastric cancer; lung cancer;

```

KW ovarian cancer; pancreatic cancer; malignant glioma;
KW uridine diphosphate glycosyltransferase1 member A1; gene; ds.

XX Homo sapiens.

XX WO2003013536-A2.

XX 20-FEB-2003.

XX 23-JUL-2002; 2002WO-EP008217.

XX 23-JUL-2001; 2001EP-00117608.

PR 24-MAY-2002; 2002EP-00011710.

XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

XX Heinrich G, Kerb R;

XX WPI; 2003-289896/28.

XX Use of irinotecan to treat cancer patient by determining if patient has
PT variant alleles of UGT1A1 gene, administering increased/decreased amounts
PT of irinotecan based on increased/decreased levels of UGT1A1 gene product.

XX Disclosure; SEQ ID NO 669; 107pp; English.

XX The invention relates to the novel use of irinotecan to treat a patient
CC suffering from cancer. This involves determining if the patient has one
CC or more variant alleles of the UGT1A1 gene, and if the patient has one or
CC more of such variant alleles, irinotecan is administered in an increased
CC or decreased amount in comparison to the amount that is administered
CC without regard to the patient's alleles in the UGT1A1 gene. The invention
CC has cytostatic activity. A composition of the invention acts as a
CC topoisomerase I inhibitor. The method is useful for treating a patient,
CC an animal e.g. mouse or a human, preferably African or Asian, suffering
CC from cancer such as colorectal, cervical, gastric cancer, lung, ovarian,
CC pancreatic cancer or malignant glioma. The present sequence is used in
CC the exemplification of the invention.

XX SQ Sequence 189013 BP; 53603 A; 38975 C; 39348 G; 57087 T; 0 U; 0 Other;

Query Match 2.0%; Score 27; DB 9; Length 189013;

Best Local Similarity 100.0%; Pred. No. 0.45;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1301 AGATGTCACAAAAAAGAAAAA 1327

Db 123937 AGATGTCACAAAAAAGAAAAA 123963

RESULT 20

ADB96928

ID ADB96928 standard; DNA; 189013 BP.

XX ADB96928;

DT 04-DEC-2003 (first entry)

XX Human MDR1 related DNA sequence SEQ ID NO:669.

XX irinotecan; colorectal cancer; cervical cancer; gastric cancer;

XX lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;

XX multidrug resistance 1; MDR1; cytostatic; human; CYP3A5; MRP1; MDR1;

XX TOPI; ds.

XX Homo sapiens.

XX WO2003013537-A2.

XX 20-FEB-2003.

XX 23-JUL-2002; 2002WO-EP008218.

XX

PR 23-JUL-2001; 2001EP-00117608.

XX 24-MAY-2002; 2002EP-00011710.

XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

XX Heinrich G, Kerb R;

XX WPI; 2003-268145/26.

XX New use of irinotecan for preparation of pharmaceutical compositions for
PT treating cancer in subject having genome with variant allele comprising
PT multidrug resistance 1 polynucleotide.

XX Disclosure; SEQ ID NO 669; 130pp; English.

XX The invention relates to the novel use of irinotecan or its derivative
CC for the preparation of pharmaceutical compositions for treating
CC colorectal, cervical, gastric, lung, ovarian or pancreatic cancer, or
CC malignant glioma in a subject having a genome with a variant allele which
CC comprises a multidrug resistance 1 (MDR1) polynucleotide. A composition
CC of the invention has cytostatic activity. The invention is useful for the
CC preparation of pharmaceutical compositions for treating colorectal
CC cancer, gastric, lung, ovarian or pancreatic cancer, or malignant
CC glioma in a subject (preferably human, more preferably African or Asian)
CC or a mouse. The present sequence is used in the exemplification of the
CC invention.

XX SQ Sequence 189013 BP; 53603 A; 38975 C; 39348 G; 57087 T; 0 U; 0 Other;

Query Match 2.0%; Score 27; DB 9; Length 189013;

Best Local Similarity 100.0%; Pred. No. 0.45;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1301 AGATGTCACAAAAAAGAAAAA 1327

Db 123937 AGATGTCACAAAAAAGAAAAA 123963

RESULT 21

ADB92119

ID ADB92119 standard; DNA; 189013 BP.

XX ADB92119;

DT 04-DEC-2003 (first entry)

XX Human MDR1 related DNA sequence SEQ ID NO:669.

XX irinotecan; colorectal cancer; cervical cancer; gastric cancer;

XX lung cancer; ovarian cancer; pancreatic cancer; malignant glioma;

XX multidrug resistance 1; MDR1; cytostatic; human; UGT1A1; MRP1; TOPI; ds.

XX Homo sapiens.

XX WO2003013535-A2.

XX 20-FEB-2003.

XX 23-JUL-2002; 2002WO-EP008220.

XX 23-JUL-2001; 2001EP-00117608.

PR 24-MAY-2002; 2002EP-00011710.

XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.

XX Heinrich G, Kerb R;

XX WPI; 2003-342400/32.

XX New use of irinotecan for preparation of pharmaceutical compositions for
PT treating cancer in subject having genome with variant allele comprising
PT multidrug resistance 1 polynucleotide.

XX

PS Disclosure; SEQ ID NO 669; 104pp; English.

XX The invention relates to a novel use of irinotecan or its derivative for

CC the preparation of a pharmaceutical composition for treating colorectal,

CC cervical, gastric, lung, ovarian or pancreatic cancer, or malignant

CC glioma in a subject having a genome with a variant allele which comprises

CC a multidrug resistance 1 (MDR1) polynucleotide. A composition of the

CC invention has cytostatic activity. The present sequence is used in the

CC exemplification of the invention.

XX

SQ Sequence 189013 BP; 53603 A; 38975 C; 39348 G; 57087 T; 0 U; 0 Other;

Query Match 2.0%; Score 27; DB 9; Length 189013;

Best Local Similarity 100.0%; Pred. No. 0.45;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

QY 1301 AGATGTCACAAAAAAGAAAAAAGAAAAA 1327

Db 123937 AGATGTCACAAAAAAGAAAAAAGAAAAA 123963

RESULT 22

ABQ60500

ID ABQ60500 standard; cDNA; 273 BP.

AC ABQ60500;

XX

XX 02-AUG-2002 (first entry)

XX Human colon cancer related nucleotide sequence SEQ ID NO:4195.

XX

XX Human; colon cancer; cancer; tissue profiling; forensic; mapping;

KW genetic analysis; diagnostic; antisense therapy; gene; ss.

XX

XX Homo sapiens.

OS

XX W0200229086-A2.

XX

XX 11-APR-2002.

PD

XX 02-OCT-2001; 2001WO-US030732.

PF

XX 02-OCT-2000; 2000US-0237271P.

XX

XX (FARB) BAYER CORP.

PA

XX Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;

PI Thielingam A, Lewis ME;

PI

XX WPI; 2002-426115/45.

DR

XX New isolated nucleic acid that is differentially expressed in cancer

PT tissues useful for determining the presence of colon cancer in a cell or

PT tissue type, and in antisense therapy.

XX

PS Claim 1; Fig 1; 796pp; English.

XX

CC ABQ63606 to ABQ60787 represent isolated nucleic acids (I) differentially

CC expressed in cancer tissues. ABB78393 to ABB79004 represent proteins

CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be

CC used in antisense therapy. An antibody immunoreactive with a polypeptide

CC encoded by (I) is useful for detecting cancer in a patient sample, and

CC for detecting the presence or absence of a polynucleotide encoded by a

CC nucleic acid which hybridizes to (I) in a cell. A probe/primer derived

CC from (I) can be used for determining the presence of a nucleic acid which

CC hybridizes to (I), and for determining the phenotype of cells in a sample

CC of cells from a patient. (I) is useful for determining the presence of

CC colon cancer in a cell or tissue type, for determining the presence or

CC state of other type of cancer, in antisense therapy, to generate

CC macroarrays on a solid surface, to identify a chromosome on which the

CC corresponding gene resides, and in tissue profiling, forensics, genetic

CC analysis, mapping and diagnostic applications. (I) can be used to raise

CC antibodies, and to screen for peptide analogues and antagonists

CC

XX

SQ Sequence 273 BP; 106 A; 29 C; 47 G; 91 T; 0 U; 0 Other;

Query Match 2.0%; Score 26; DB 6; Length 273;

Best Local Similarity 100.0%; Pred. No. 3.3;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX

QY 1302 GATGTCACAAAAAAGAAAAAAGAAAAA 1327

Db 245 GATGTCACAAAAAAGAAAAAAGAAAAA 270

RESULT 23

ABX41090/c

ID ABX41090 standard; cDNA; 393 BP.

XX

XX ABX41090;

AC

XX 20-FEB-2003 (first entry)

XX

XX Bovine EST associated with lactation/muscle/fat deposition #6255.

DE

XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;

KW muscle deposition; fat deposition; genome mapping; gene identification;

KW gene analysis; cattle breeding.

XX

XX Bos Taurus.

OS

XX US2002137139-A1.

XX

XX 26-SEP-2002.

PD

XX 24-SEP-2001; 2001US-00960352.

PF

XX 12-JAN-1999; 99US-0115707P.

PR

XX 11-JAN-2000; 2000US-00480902.

XX

XX (BYAT/) BYATT J C.

PA (MATH/) MATHIALAGAN N.

PA (TAON/) TAO N.

PA (WARR/) WARREN W C.

XX

XX Byatt JC, Mathialagan N, Tao N, Warren WC;

PI

XX WPI; 2003-110599/10.

DR

XX New nucleic acid associated with lactation, and muscle and fat

PT deposition, useful for genome mapping, gene identification and analysis,

PT cattle breeding, or for genetically improving cattle.

XX

PS Claim 2; SEQ ID NO 6255; 245pp; English.

XX

CC The invention relates to a purified nucleic acid molecule associated with

CC lactation or muscle and fat deposition (designated LMFD), derived from

CC cattle, and the LMFD nucleic acid can specifically hybridize to a second

CC nucleic acid molecule comprising any of 15112 nucleotide sequences,

CC appearing as ABX34838-ABX49947, or complements of them. Also included are

CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic

CC acid linked to a promoter and a 3' non-translated sequence that

CC functions in the cell to cause termination of transcription and addition

CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and

CC (2) determining a level or pattern of a molecule in a bovine cell or

CC tissue comprising: (a) incubating a marker nucleic acid (comprising any

CC of the 15112 nucleic acid sequences or its complement or fragment) with a

CC complementary nucleic acid molecule obtained from the bovine cell or

CC tissue, where hybridisation between the marker nucleic acid and the

CC complementary nucleic acid permits the detection of the molecule; and (b)

CC detecting the level or pattern of the complementary nucleic acid, where

CC the detection of the complementary nucleic acid is predictive of the

CC level or pattern of the molecule. The LMFD nucleic acid is used for

CC determining a level or pattern of a molecule in a bovine cell or tissue.

CC It is useful for genome mapping, gene identification and analysis, cattle

CC breeding, preparation of constructs for use in cattle gene expression, or

CC

CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139
 XX
 SQ Sequence 393 BP; 107 A; 75 C; 96 G; 115 T; 0 U; 0 Other;
 Query Match 2.0%; Score 26; DB 7; Length 393;
 Best Local Similarity 100.0%; Pred. No. 3.1;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 SQ
 QY 1302 GATGTCACAAAAA 1327
 |||||
 Db 49 GATGTCACAAAAA 24
 |||||
 RESULT 24
 AX48150/c
 ID AX48150 standard; cDNA; 411 BP.
 XX
 AC AX48150;
 XX
 XX
 DT 21-FEB-2003 (first entry)
 XX
 DE Bovine EST associated with lactation/muscle/fat deposition #13315.
 XX
 KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 KW muscle deposition; fat deposition; genome mapping; gene identification;
 KW gene analysis; cattle breeding.
 XX
 XX Bos Taurus.
 OS
 XX
 FN US2002137139-A1.
 XX
 XX
 PD 26-SEP-2002.
 XX
 PF 24-SEP-2001; 2001US-00960352.
 XX
 PR 12-JAN-1999; 99US-0115707P.
 PR 11-JAN-2000; 2000US-00480902.
 XX
 XX (BYAT/) BYATT J C.
 PA (MATH/) MATHIALAGAN N.
 PA (TAON/) TAO N.
 PA (WARR/) WARREN W C.
 XX
 PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX
 DR WPI; 2003-110599/10.
 XX
 XX New nucleic acid associated with lactation, and muscle and fat
 PT deposition, useful for genome mapping, gene identification and analysis,
 PT cattle breeding, or for genetically improving cattle.
 XX
 PS Claim 2; SEQ ID NO 13315; 245pp; English.
 XX
 XX The invention relates to a purified nucleic acid molecule associated with
 CC lactation or muscle and fat deposition (designated LMFD), derived from
 CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
 CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
 CC appearing as ABX34836-ABX49947, or complements of them. Also included are
 CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
 CC acid linked to a promoter and a 3' non- translated sequence that
 CC functions in the cell to cause termination of transcription and addition
 CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 CC (2) determining a level or pattern of a molecule in a bovine cell or
 CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
 CC of the 15112 nucleic acid sequences or its complement or fragment) with a
 CC complementary nucleic acid molecule obtained from the bovine cell or
 CC tissue, where hybridisation between the marker nucleic acid and the
 CC complementary nucleic acid permits the detection of the molecule; and (b)
 CC detecting the level or pattern of the complementary nucleic acid, where

CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification but was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139
 XX
 SQ Sequence 411 BP; 110 A; 80 C; 99 G; 122 T; 0 U; 0 Other;
 Query Match 2.0%; Score 26; DB 7; Length 411;
 Best Local Similarity 100.0%; Pred. No. 3.1;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 SQ
 QY 1302 GATGTCACAAAAA 1327
 |||||
 Db 48 GATGTCACAAAAA 23
 |||||
 RESULT 25
 AAL24218/c
 ID AAL24218 standard; cDNA; 659 BP.
 XX
 AC AAL24218;
 XX
 DT 07-DEC-2001 (first entry)
 XX
 DE Human breast cancer expressed polynucleotide 16675.
 XX
 KW Human; breast cancer; cell marker; cytostatic; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200151628-A2.
 XX
 PD 19-JUL-2001.
 XX
 PF 10-JAN-2001; 2001WO-US000798.
 XX
 PR 14-JAN-2000; 2000US-0176077P.
 PR 14-MAR-2000; 2000US-0189167P.
 PR 24-MAR-2000; 2000US-0192099P.
 PR 29-MAR-2000; 2000US-0193480P.
 PR 15-MAY-2000; 2000US-0205230P.
 PR 09-JUN-2000; 2000US-0211315P.
 PR 25-JUL-2000; 2000US-0220534P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Lillie J, Xu Y, Wang Y, Steinmann K;
 XX
 DR WPI; 2001-451856/48.
 XX
 PT New peptide useful as a marker for the diagnosis of breast cancer.
 XX
 PS Claim 1; Page 3055; 3695pp; English.
 XX
 XX The invention relates to human breast cancer expressed polynucleotides
 CC (AAL07544-AAL26789) and methods of assessing whether a patient is
 CC afflicted with breast cancer by examining the correlation between the
 CC expression of certain markers and the cancerous state of breast cells.
 CC The polynucleotides and encoded polypeptides are potential markers for
 CC detecting, diagnosing, monitoring, characterising treating and
 CC potentially preventing breast cancer. The polynucleotides and encoded
 CC polypeptides are also useful for isolating compounds with cytostatic
 CC activity
 XX
 SQ Sequence 659 BP; 184 A; 181 C; 110 G; 181 T; 0 U; 3 Other;
 Query Match 2.0%; Score 26; DB 4; Length 659;

Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1302 GATGTCACAAAAA 1327
Db 599 GATGTCACAAAAA 574

RESULT 26
AAQ35208
ID AAQ35208 standard; cDNA; 729 BP.
XX AC AAQ35208;
XX
DT 06-JUN-1993 (first entry)
XX
DE Sequence encoding adenine phosphoribosyltransferase (apt).
XX
KW Adenine phosphoribosyltransferase; probe; APRT activity; ss.
XX
OS Arabidopsis thaliana.
XX
FH Key Location/Qualifiers
FT CDS 18..569
FT /*tag= a

XX CA2069262-A.
XX
XX 24-NOV-1992.
XX
XX 22-MAY-1992; 92CA-02069262.
XX
XX 23-MAY-1991; 91GB-000111126.
XX
XX (UYWA-) UNIV WATERLOO.
XX
XX Moffatt B;
XX
XX WPI; 1993-053310/07.
XX P-PSDB; AAR30826.
XX
XX Purified c DNA for apt. gene in plants - used for analysis of adenine
XX phosphoribosyl-transferase activity and function in plant development.
XX
XX Claim 3; Fig 1; 27pp; English.
XX
XX The apt cDNA of A. thaliana is described as a full-length clone. However,
XX the inventors cannot unequivocally rule out that there are 5' sequences
XX missing from this cDNA. The cDNA predicts a protein molecular weight of
XX 27,140
XX
SQ Sequence 729 BP; 204 A; 123 C; 188 G; 214 T; 0 U; 0 Other;

Query Match 2.0%; Score 26; DB 2; Length 729;
Best Local Similarity 100.0%; Pred. No. 2.8;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1302 GATGTCACAAAAA 1327
Db 703 GATGTCACAAAAA 728

RESULT 27
AAS02396
ID AAS02396 standard; cDNA; 2295 BP.
XX AC AAS02396;
XX
DT 18-JUL-2001 (first entry)
XX
DE Human secreted protein, cDNA #2.
XX
KW Human; secreted protein; immunogen; antibody; diagnosis;

KW rheumatoid arthritis; hyperproliferative disorder; neoplasm;
KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;
KW angiogenesis; Alzheimer's disease; bacterial infection; viral infection;
KW fungal infection; corneal infection; wound healing; cell culture;
KW epithelial cell proliferation; skin ageing; transplantation;
KW tissue regeneration; chemotaxis; food additive; ss.
XX
OS Homo sapiens.
XX
XX WO200123546-A1.
XX
XX 05-APR-2001.
XX
XX 26-SEP-2000; 2000WO-US026323.
XX
XX 27-SEP-1999; 99US-0155805P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben S, Komatsoulis GA;
XX
XX WPI; 2001-266150/27.
XX P-PSDB; AAU01723, AAU01759, AAU01760.
XX
XX Nucleic acids encoding 37 human secreted polypeptides, useful for
XX preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease
XX and diabetic retinopathy.
XX
XX Disclosure; Page 411; 494pp; English.
XX
XX The sequence encodes a human secreted protein of the invention. The
XX polynucleotides, polypeptides and antibodies raised against them are used
XX to prevent, treat or ameliorate a medical condition in e.g. humans, mice,
XX rabbits, goats, horses, cats, dogs, chickens or sheep. The
XX polynucleotides and antibodies are also used in diagnosing a pathological
XX condition or susceptibility to a pathological condition. The antibodies
XX can also be used in alleviating symptoms associated with the disorders
XX and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked
XX immunosorbent assays (ELISA). Disorders which are diagnosed or treated
XX include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
XX disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
XX e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
XX angiogenesis, nervous system disorders e.g. Alzheimer's disease,
XX infections caused by bacteria, viruses and fungi and ocular disorders
XX e.g. corneal infection. The polypeptides can also be used to aid wound
XX healing and epithelial cell proliferation, to prevent skin aging due to
XX sunburn, to maintain organs before transplantation, for supporting cell
XX culture of primary tissues, to regenerate tissues and in chemotaxis. The
XX polypeptides can also be used as a food additive or preservative to
XX increase or decrease storage capabilities
XX
SQ Sequence 2295 BP; 794 A; 393 C; 449 G; 659 T; 0 U; 0 Other;

Query Match 2.0%; Score 26; DB 4; Length 2295;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1301 AGATGTCACAAAAA 1326
Db 2270 AGATGTCACAAAAA 2295

RESULT 28
ABX76356
ID ABX76356 standard; DNA; 4702 BP.
XX AC ABX76356;
XX
XX 02-APR-2003 (first entry)
XX
DE Lung cancer-associated polynucleotide #220.
XX
KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;


```

XX SQ Sequence 184 BP; 76 A; 30 C; 50 G; 28 T; 0 U; 0 Other;
Query Match      1.9%; Score 25; DB 6; Length 184;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCACAAAAA...AAAAAAAAA 1327
Db 141 ATGTCACAAAAA...AAAAAAAAA 165

RESULT 31
ID ABL65710/c
XX ABL65710 standard; DNA; 191 BP.
AC ABL65710;
XX
DT 15-MAY-2002 (first entry)
XX
DE Lung cancer related gene sequence SEQ ID NO:4047.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.
PR 18-SEP-2000; 2000US-0233617P.
PR 20-SEP-2000; 2000US-0234009P.
PR 20-SEP-2000; 2000US-0234034P.
PR 20-SEP-2000; 2000US-0234052P.
PR 22-SEP-2000; 2000US-0234509P.
PR 22-SEP-2000; 2000US-0234567P.
PR 25-SEP-2000; 2000US-0234923P.
PR 25-SEP-2000; 2000US-0234924P.
PR 25-SEP-2000; 2000US-0235077P.
PR 25-SEP-2000; 2000US-0235082P.
PR 25-SEP-2000; 2000US-0235134P.
PR 25-SEP-2000; 2000US-0235280P.
PR 26-SEP-2000; 2000US-0235637P.
PR 26-SEP-2000; 2000US-0235638P.
PR 27-SEP-2000; 2000US-0235711P.
PR 27-SEP-2000; 2000US-0235720P.
PR 27-SEP-2000; 2000US-0235840P.
PR 27-SEP-2000; 2000US-0235863P.
PR 28-SEP-2000; 2000US-0236028P.
PR 28-SEP-2000; 2000US-0236032P.
PR 28-SEP-2000; 2000US-0236033P.
PR 28-SEP-2000; 2000US-0236034P.
PR 28-SEP-2000; 2000US-0236109P.
PR 28-SEP-2000; 2000US-0236111P.
PR 29-SEP-2000; 2000US-0236842P.
PR 29-SEP-2000; 2000US-0236891P.
PR 02-OCT-2000; 2000US-0237172P.
PR 02-OCT-2000; 2000US-0237173P.
PR 02-OCT-2000; 2000US-0237278P.
PR 02-OCT-2000; 2000US-0237294P.
PR 02-OCT-2000; 2000US-0237295P.
PR 02-OCT-2000; 2000US-0237316P.
PR 03-OCT-2000; 2000US-0237425P.
PR 03-OCT-2000; 2000US-0237598P.
PR 03-OCT-2000; 2000US-0237604P.

PR 03-OCT-2000; 2000US-0237606P.
PR 03-OCT-2000; 2000US-0237608P.
PR 01-NOV-2000; 2000US-0244867P.
PR 01-NOV-2000; 2000US-0245084P.
XX (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, and determining a change
XX in expression of a gene of a signature gene set.
XX
XX Claim 1; SEQ ID NO 4047; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an anti-
XX neoplastic agent. The method involves exposing cells to a chemical agent
XX to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic
XX activity and can be used in gene therapy. M1 can be used for screening an
XX anti-neoplastic agent, and can be used for producing a product which is
XX the data collected with respect to the anti-neoplastic agent as a result
XX of M1, and the data is sufficient to convey the chemical structure and/or
XX properties of the agent. M1 can be used in the treatment of cancer such
XX as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,
XX prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
XX cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
XX cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's
XX tumour
XX
XX SQ Sequence 191 BP; 52 A; 38 C; 23 G; 78 T; 0 U; 0 Other;
Query Match      1.9%; Score 25; DB 6; Length 191;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCACAAAAA...AAAAAAAAA 1327
Db 30 ATGTCACAAAAA...AAAAAAAAA 6

RESULT 32
ABL64256/c
ID ABL64256 standard; DNA; 191 BP.
XX
AC ABL64256;
XX
DT 15-MAY-2002 (first entry)
XX
DE Stomach cancer related gene sequence SEQ ID NO:2593.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US010838.
XX
PR 05-JUN-2000; 2000US-0209473P.
PR 05-JUN-2000; 2000US-0209531P.
PR 18-SEP-2000; 2000US-0233133P.

```

18-SEP-2000; 2000US-0233617P.
20-SEP-2000; 2000US-0234009P.
20-SEP-2000; 2000US-0234034P.
20-SEP-2000; 2000US-0234052P.
22-SEP-2000; 2000US-0234509P.
22-SEP-2000; 2000US-0234567P.
23-SEP-2000; 2000US-0234923P.
25-SEP-2000; 2000US-0234924P.
25-SEP-2000; 2000US-0235077P.
25-SEP-2000; 2000US-0235082P.
25-SEP-2000; 2000US-0235134P.
25-SEP-2000; 2000US-0235280P.
26-SEP-2000; 2000US-0235637P.
26-SEP-2000; 2000US-0235638P.
27-SEP-2000; 2000US-0235711P.
27-SEP-2000; 2000US-0235720P.
27-SEP-2000; 2000US-0235840P.
27-SEP-2000; 2000US-0235863P.
28-SEP-2000; 2000US-0236032P.
28-SEP-2000; 2000US-0236033P.
28-SEP-2000; 2000US-0236034P.
28-SEP-2000; 2000US-0236109P.
28-SEP-2000; 2000US-0236111P.
29-SEP-2000; 2000US-0236842P.
29-SEP-2000; 2000US-0236891P.
02-OCT-2000; 2000US-0237172P.
02-OCT-2000; 2000US-0237173P.
02-OCT-2000; 2000US-0237278P.
02-OCT-2000; 2000US-0237294P.
02-OCT-2000; 2000US-0237295P.
02-OCT-2000; 2000US-0237316P.
03-OCT-2000; 2000US-0237425P.
03-OCT-2000; 2000US-0237598P.
03-OCT-2000; 2000US-0237604P.
03-OCT-2000; 2000US-0237606P.
03-OCT-2000; 2000US-0237608P.
01-NOV-2000; 2000US-0244867P.
01-NOV-2000; 2000US-0245084P.
(AVAL-) AVALON PHARM.
Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
Soppet DR, Weaver Z;
WPI; 2002-188264/24.
Screening for anti-neoplastic agent involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, and determining a change in expression of a gene of a signature gene set.
Claim 1; SEQ ID NO 2593; 44pp; English.
The present invention describes a method (M1) for screening for an anti-neoplastic agent. The method involves exposing cells to a chemical agent to be tested for anti-neoplastic activity, determining a change in expression of at least one gene (I) of a signature gene set, where (I) comprises a sequence (S) selected from 8447 sequences (given in ABL61664 to ABL70110), or is at least 95% identical to (S), where a change in expression is indicative of anti-neoplastic activity. (I) has cytostatic activity and can be used in gene therapy. M1 can be used for screening an anti-neoplastic agent, and can be used for producing a product which is the data collected with respect to the anti-neoplastic agent as a result of M1, and the data is sufficient to convey the chemical structure and/or properties of the agent. M1 can be used in the treatment of cancer such as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney, prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's tumour
Sequence 191 BP; 52 A; 38 C; 23 G; 78 T; 0 U; 0 Other;

Query Match 1.9%; Score 25; DB 6; Length 191;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1303 ATGTCACAAAAA 1327
DB 30 ATGTCACAAAAA 6
RESULT 33
ABN94027/c
ID ABN94027 standard; DNA; 191 BP.
XX AC ABN94027;
XX 13-AUG-2002 (first entry)
XX Gene #525 used to diagnose liver cancer.
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX Homo sapiens.
XX WO200229103-A2.
XX 11-APR-2002.
XX 02-OCT-2001; 2001WO-US030589.
XX 02-OCT-2000; 2000US-0237054P.
XX (GENE-) GENE LOGIC INC.
XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX WPI; 2002-426119/45.
Diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumor in a patient, involves detecting the level of expression of two or more genes in a liver tissue sample.
Claim 1; SEQ ID NO 525; 298pp; English.
The invention relates to a novel method for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma or metastatic liver tumor in a patient, and differentiating metastatic liver cancer from hepatocellular carcinoma in a patient, involving detecting the level of expression of two or more genes represented in ABN93503-ABN97455 in a tissue sample. The method of the invention has hepatotropic, and cytostatic activity. The method is useful for diagnosing and detecting the progression of liver cancer, hepatocellular carcinoma and metastatic liver carcinoma in a patient. The method is useful for identifying expression profiles which serve as useful diagnostic markers as well as markers that can be used to monitor disease states, disease progression, drug toxicity, drug efficacy and drug metabolism. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
Sequence 191 BP; 52 A; 38 C; 23 G; 78 T; 0 U; 0 Other;
Query Match 1.9%; Score 25; DB 6; Length 191;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1303 ATGTCACAAAAA 1327
DB 30 ATGTCACAAAAA 6
RESULT 34

XX	ABL87801	Gene #587 used to diagnose liver cancer.
DE	XX	
XX	XX	
XX	XX	
KW	Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;	
AC	metastatic liver tumour; cytostatic; expression profile; disease state;	
XX	metastatic liver tumour; cytostatic; expression profile; disease state;	
DT	17-MAY-2002 (first entry)	
XX	Human ovarian cancer related cDNA clone SEQ ID NO:10779.	
XX	Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.	
DE	Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.	
KW	Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.	
XX	Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.	
OS	Homo sapiens.	
XX	Homo sapiens.	
XX	WO200192581-A2.	
EN	06-DEC-2001.	
XX	29-MAY-2001; 2001WO-US017756.	
XX	26-MAY-2000; 2000US-0207484P.	
PF	(CORI-) CORIXA CORP.	
XX	Algate PA, Harlocker SL, Jones R;	
XX	WPI; 2002-122075/16.	
DR	Composition for therapy and diagnosis of ovarian cancer comprising	
PT	polypeptide of a ovarian tumor polypeptide, polynucleotide encoding	
PT	polypeptide, antibody specific to polypeptide or T cell expressing	
PT	polypeptide.	
XX	Claim 1; SEQ ID NO 10779; 489pp; English.	
PS	The present invention describes a composition (I) comprising: carriers	
XX	and immunostimulants; and a polypeptide (II) of a ovarian tumour	
CC	polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)	
CC	from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,	
CC	(III) encoding (II) having a sequence (S2), a T cell population of (II),	
CC	or antigen presenting cells that express (II). (I) has cytostatic	
CC	activity. An oligonucleotide (IV) that hybridises to (S1) can be used for	
CC	detecting ovarian cancer in a patient's biological sample preferably	
CC	serum or ovarian tissue. The method comprises contacting a biological	
CC	sample from a patient with (IV), detecting the amount of polynucleotide	
CC	hybridising to (IV) and comparing the amount to a predetermined cutoff	
CC	value and thereby detecting ovarian cancer in the patient, where the	
CC	amount of polynucleotide hybridising to (IV) is detected preferably by	
CC	polymerase chain reaction (PCR). (I) comprising (III) and/or (II) is	
CC	useful for stimulating and/or expanding T cells specific for an ovarian	
CC	tumour protein comprising contacting T cells with (III) or (II). (III) is	
CC	useful in design and preparation of ribozyme molecules for inhibiting	
CC	expression of the tumour polypeptides and proteins in tumour cells; and	
CC	to isolate a full length gene from a suitable library e.g., a tumour cDNA	
CC	library using well known techniques	
XX	Sequence 207 BP; 81 A; 38 C; 54 G; 34 T; 0 U; 0 Other;	
SQ		
Query Match	1.9%; Score 25; DB 6; Length 207;	
Best Local Similarity	100.0%; Pred. No. 8.4;	
Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1303 ATGTCACAAAAAATAAAAAAAAAA 1327	
DB	164 ATGTCACAAAAAATAAAAAAAAAA 188	
RESULT 35		
ABN94089/c		
ID	ABN94089 standard; DNA; 222 BP.	
XX	ABN94089;	
AC		
XX	13-AUG-2002 (first entry)	
DT		

```
PD 27-SEP-2001.
XX
XX
XX 20-MAR-2001; 2001WO-US0009062.
XX
XX 21-MAR-2000; 2000US-0190710P.
PR 22-JUN-2000; 2000US-0213748P.
PR 19-DEC-2000; 2000US-0257276P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Xu J, Pyle RA, Stolk JA;
XX WPI; 2001-607531/69.
XX
XX Nucleic acids encoding 222 polypeptides associated with ovarian and
PT endometrial cancers, useful for diagnosing, preventing and treating
PT cancers.
XX
XX Claim 1; Page 137-138; 187pp; English.
XX
XX The invention relates to human polynucleotides encoding proteins
CC associated with ovarian and endometrial cancers. The polynucleotides and
CC the proteins they encode may be used in the prevention, diagnosis and
CC treatment of diseases associated with the inappropriate expression of
CC ovarian and endometrial cancer polypeptides (OECs). For example, the
CC polynucleotide (or an expression vector comprising the polynucleotide)
CC and the OECs may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of OECs by expressing inactive proteins or to
CC supplement the patient's own production of them. Additionally, the
CC polynucleotide may be used to produce the OECs, by inserting the nucleic
CC acids into a host cell and culturing the cell to express the protein. The
CC polynucleotide and its complementary sequences may also be used as DNA
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acids in samples, and therefore which patients may be in
CC need of restorative therapy. The OECs may also be used as antigens in
CC the production of anti-OEC antibodies and in assays to identify
CC modulators of its expression and activity. The anti-OEC antibodies and
CC antagonists may also be used to down regulate expression and activity.
CC The anti-OEC antibodies may also be used as diagnostic agents for
CC detecting the presence of OEC in samples (e.g. by enzyme linked
CC immunosorbent assay (ELISA)) and hence diagnose patients with cancers.
CC The present sequence is a ovarian and endometrial cancer linked cDNA of
CC the invention
XX
XX Sequence 330 BP; 83 A; 76 C; 64 G; 103 T; 0 U; 4 Other;
SQ
Query Match 1.9%; Score 25; DB 4; Length 330;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1303 ATGTCACAAAAA 1327
Db 30 ATGTCACAAAAA 6
RESULT 37
ID ABX36291/c
XX ABX36291 standard; cDNA; 354 BP.
XX
XX ABX36291;
XX
XX 20-FEB-2003 (first entry)
XX
XX Bovine EST associated with lactation/muscle/fat deposition #1456.
XX
XX Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
XX Bos Taurus.
XX
XX OS
XX
XX US2002137139-A1.
XX
```

```
XX
XX
XX 26-SEP-2002.
XX
XX 24-SEP-2001; 2001US-00960352.
XX
XX 12-JAN-1999; 99US-0115707P.
PR 11-JAN-2000; 2000US-00480902.
XX
XX (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARR/) WARREN W C.
XX
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-110599/10.
XX
XX New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX
XX Claim 2; SEQ ID NO 1456; 245pp; English.
XX
XX The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (Designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridize to a second
CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non-translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 15112 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridisation between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and (b)
CC detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMFD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?docID=20020137139
XX
XX Sequence 354 BP; 92 A; 65 C; 92 G; 105 T; 0 U; 0 Other;
SQ
Query Match 1.9%; Score 25; DB 7; Length 354;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1302 GATGTCACAAAAA 1326
Db 39 GATGTCACAAAAA 15
RESULT 38
ID AAL09871/c
XX AAL09871 standard; cDNA; 365 BP.
XX
XX AAL09871;
XX
XX 07-DEC-2001 (first entry)
XX
XX Human breast cancer expressed polynucleotide 2328.
XX
XX Human; breast cancer; cell marker; cytostatic; ss.
XX
```

```

OS Homo sapiens.
XX WO200151628-A2.
XX 19-JUL-2001.
XX
XX 10-JAN-2001; 2001WO-US000798.
XX
XX 14-JAN-2000; 2000US-0176077P.
XX 14-MAR-2000; 2000US-0189167P.
XX 24-MAR-2000; 2000US-0192099P.
XX 29-MAR-2000; 2000US-0193480P.
XX 15-MAY-2000; 2000US-0205230P.
XX 09-JUN-2000; 2000US-0211315P.
XX 25-JUL-2000; 2000US-0220534P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Lillie J, Xu Y, Wang Y, Steinmann K;
XX WPI; 2001-451856/48.
XX
XX New peptide useful as a marker for the diagnosis of breast cancer.
XX
XX Claim 1; Page 446; 3695pp; English.
XX
XX The invention relates to human breast cancer expressed polynucleotides
XX (AAL07544-AAL26789) and methods of assessing whether a patient is
XX afflicted with breast cancer by examining the correlation between the
XX expression of certain markers and the cancerous state of breast cells.
XX The polynucleotides and encoded polypeptides are potential markers for
XX detecting, diagnosing, monitoring, characterising treating and
XX potentially preventing breast cancer. The polynucleotides and encoded
XX polypeptides are also useful for isolating compounds with cytostatic
XX activity
XX
XX Sequence 365 BP; 99 A; 78 C; 65 G; 111 T; 0 U; 12 Other;
XX
XX Query Match 1.9%; Score 25; DB 4; Length 365;
XX Best Local Similarity 100.0%; Pred. No. 7.6;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1303 ATGTCAAAAA AAAAAAAAAA 1327
XX Db 86 ATGTCAAAAA AAAAAAAAAA 62
XX
XX RESULT 39
XX ACC48717
XX ID ACC48717 standard; cDNA; 390 BP.
XX
XX AC ACC48717;
XX
XX DT 11-AUG-2003 (first entry)
XX
XX DE Mouse wound healing biomarker EST AA608387 103723_at.
XX
XX KW Mouse; wound healing; biomarker; marker; vulnerary; vasotropic;
XX antipsoriatic; cytostatic; ophthalmological; antiulcer; gene therapy;
XX expressed sequence tag; EST; ss.
XX
XX OS Mus musculus.
XX
XX PN BP1270044-A2.
XX
XX PD 02-JAN-2003.
XX
XX 17-JUN-2002; 2002EP-00254207.
XX
XX 18-JUN-2001; 2001GB-00014869.
XX 13-JUL-2001; 2001US-0305346P.
XX
XX (PFIZ ) PFIZER LTD.

```

```

PA (PFIZ ) PFIZER INC.
XX
XX Burslem MF, Cooper L, Johnson CM, Martin P;
XX WPI; 2003-291795/29.
XX
XX Novel wound healing biomarkers including genes and protein products
XX identified as being involved in wound healing purposes, useful for
XX identifying compounds for treating wounds, wound healing disorders or
XX inflammation.
XX
XX Claim 1; Page 165; 176pp; English.
XX
XX The present DNA sequence corresponds to an expressed sequence tag (EST)
XX identified from an analysis of murine genes that are up- or down-
XX regulated during a wound healing response, with or without the
XX inflammatory response usually occurring in healing wounds. The gene that
XX encodes this EST encodes a protein important in the late, advanced phase
XX of wound healing, i.e. in the cessation of the healing response,
XX maturation and remodelling of the repaired tissue. A gene comprising this
XX EST, a protein encoded by the gene, the human homologue of the gene, and
XX a biomarker derived from the gene are claimed. Antagonists and inhibitors
XX of this novel target (e.g. antibodies, antisense oligonucleotides and
XX ribozymes) can be used to treat wounds or disorders characterised by
XX excessive wound healing, such as scarring, fibrosis, restenosis post-
XX angioplasty, post-traumatic/surgical adhesions of the peritoneal cavity,
XX joints and ligaments, psoriasis, benign prostatic hyperplasia, glaucoma
XX or peripheral nerve injury. Agonists or activators of this target (e.g.
XX the target protein delivered by means of gene therapy) can be used to
XX treat a disease or disorder characterised by impaired healing response,
XX such as chronic dermal ulcers, oral mucocystitis, emphysema, ulcerative
XX diseases of the gastrointestinal tract and cystitis. Methods are provided
XX for monitoring the progress of wound healing and for identification of
XX individuals with wound healing disorders
XX
XX Sequence 390 BP; 143 A; 59 C; 72 G; 116 T; 0 U; 0 Other;
XX
XX Query Match 1.9%; Score 25; DB 7; Length 390;
XX Best Local Similarity 100.0%; Pred. No. 7.5;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1303 ATGTCAAAAA AAAAAAAAAA 1327
XX Db 356 ATGTCAAAAA AAAAAAAAAA 380
XX
XX RESULT 40
XX AAL18543/c
XX ID AAL18543 standard; cDNA; 414 BP.
XX
XX AC AAL18543;
XX
XX DT 07-DEC-2001 (first entry)
XX
XX DE Human breast cancer expressed polynucleotide 11000.
XX
XX KW Human; breast cancer; cell marker; cytostatic; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200151628-A2.
XX
XX PD 19-JUL-2001.
XX
XX 10-JAN-2001; 2001WO-US000798.
XX
XX 14-JAN-2000; 2000US-0176077P.
XX 14-MAR-2000; 2000US-0189167P.
XX 24-MAR-2000; 2000US-0192099P.
XX 29-MAR-2000; 2000US-0193480P.
XX 15-MAY-2000; 2000US-0205230P.
XX 09-JUN-2000; 2000US-0211315P.
XX 25-JUL-2000; 2000US-0220534P.
XX

```

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX Lillie J, Xu Y, Wang Y, Steinmann K;
 XX WPI; 2001-451856/48.
 XX New peptide useful as a marker for the diagnosis of breast cancer.
 XX Claim 1; Page 1959; 3695pp; English.
 XX The invention relates to human breast cancer expressed polynucleotides
 CC (AAL07544-AAL26789) and methods of assessing whether a patient is
 CC afflicted with breast cancer by examining the correlation between the
 CC expression of certain markers and the cancerous state of breast cells.
 CC The polynucleotides and encoded polypeptides are potential markers for
 CC detecting, diagnosing, monitoring, characterizing treating and
 CC potentially preventing breast cancer. The polynucleotides and encoded
 CC polypeptides are also useful for isolating compounds with cytostatic
 CC activity
 XX Sequence 414 BP; 117 A; 83 C; 73 G; 140 T; 0 U; 1 Other;
 SQ
 Query Match 1.9%; Score 25; DB 4; Length 414;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1303 ATGTCACAAAAAAGAAAAAAGAAAAA 1327
 Dd 47 ATGTCACAAAAAAGAAAAAAGAAAAA 23
 RESULT 41
 ABV18188
 ID ABV18188 standard; cDNA; 425 BP.
 XX
 AC ABV18188;
 XX
 DT 13-SEP-2002 (first entry)
 DE Human prostate expression marker cDNA 18179.
 XX
 XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; Gene; ss.
 XX
 OS Homo sapiens.
 XX WO200160860-A2.
 FN 23-AUG-2001.
 PD 20-FEB-2001; 2001WO-US005171.
 PF 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX Claim 1; Page 3000; 11750pp; English.
 PS The invention relates to an isolated nucleic acid molecule (i) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (i) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 425 BP; 185 A; 75 C; 40 G; 89 T; 0 U; 36 Other;
 Query Match 1.9%; Score 25; DB 5; Length 425;
 Best Local Similarity 100.0%; Pred. No. 7.4;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1303 ATGTCACAAAAAAGAAAAAAGAAAAA 1327
 Dd 131 ATGTCACAAAAAAGAAAAAAGAAAAA 155
 RESULT 42
 ACH20487
 ID ACH20487 standard; cDNA; 430 BP.
 XX
 AC ACH20487;
 XX
 DT 13-OCT-2003 (first entry)
 DE Human adult liver cDNA #99.
 XX
 XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
 KW genome mapping; biodiversity; genetic disorder.
 XX
 OS Homo sapiens.
 XX US2003073623-A1.
 FN 17-APR-2003.
 PD 30-JUL-2001; 2001US-00918995.
 PF 30-JUL-2001; 2001US-00918995.
 PR (DRNA/) DRMANAC R T.
 PA (LABA/) LABAT I.
 PA (STAC/) STACHE-CRAIN B.
 PA (DICK/) DICKSON M C.
 PA (JONE/) JONES L W.
 XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
 PI WPI; 2003-615964/58.
 XX New polynucleotide sequences obtained from various cDNA libraries, useful
 PT as hybridization probes, as oligomers for PCR, for chromosome and gene
 PT mapping, in the recombinant production of protein, or in generating
 PT antisense DNA or RNA.
 XX Claim 1; SEQ ID NO 7699; 44pp; English.
 XX The invention relates to an isolated polynucleotide comprising any one of
 CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
 CC determined by the technique of SBH (sequencing by hybridisation). Also
 CC included is a purified polypeptide comprising a sequence corresponding to
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences
 CC are useful in diagnostics as expressed sequence tags (EST) for
 CC identifying expressed genes or for physical mapping of the human genome,
 CC in forensics, in assessing biodiversity, or in identifying mutations
 CC responsible for genetic disorders and other traits. The nucleotide
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,

CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ Sequence 430 BP; 117 A; 110 C; 100 G; 103 T; 0 U; 0 Other;

Query Match 1.9%; Score 25; DB 8; Length 430;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 405 ATGTCACAAAAA 429

RESULT 43
ACF03901
ID ACF03901 standard; cDNA; 433 BP.
AC ACF03901;
XX
DT 19-SEP-2003 (first entry)
XX
DE Rice endosperm expressed sequence tag (EST) cDNA SEQ ID NO:6.
XX
DE Rice; plant; endosperm; expressed sequence tag; EST; gene chip;
KW expression sequence label; gene; ss.
XX
OS Oryza sativa.
XX
PN CN1366069-A.
XX
PD 28-AUG-2002.
XX
PF 31-OCT-2001; 2001CN-00135857.
PR 31-OCT-2001; 2001CN-00135857.
XX
PA (UYZH-) UNIV ZHEJIANG.
XX
PI Dong H, Li D;
XX
DR WPI; 2003-392546/37.
XX
PT Rice endosperm specific expression sequence label and gene chip prepared
PT by it.
XX
PS Claim 1; Page 12 (Disclosure); 28pp; Chinese.
XX
CC The present invention describes the rice endosperm specific expressed
CC sequence tag (EST) cDNA sequences given in ACF03896 to ACF03943 (SEQ ID
CC NO:1 to 50, where SEQ ID NO:16 and 44 have not been given in the
CC as specification), isolated from a rice cDNA library. The ESTs are described
CC as expression sequence labels. Also described are gene chips constituted
CC from the expression sequence labels. The expression sequence labels can
CC be used for DNA sequencing to a configured rice endosperm cDNA library
XX
SQ Sequence 433 BP; 107 A; 93 C; 104 G; 129 T; 0 U; 0 Other;

Query Match 1.9%; Score 25; DB 7; Length 433;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 409 ATGTCACAAAAA 433

RESULT 44

ABX43422/c
ID ABX43422 standard; cDNA; 464 BP.
XX
AC ABX43422;
XX
DT 21-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #8587.
XX
KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX
PN US2002137139-A1.
XX
PD 26-SEP-2002.
XX
PF 24-SEP-2001; 2001US-00960352.
XX
PR 12-JAN-1999; 99US-0115707P.
PR 11-JAN-2000; 2000US-00480902.
XX
PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARRE/) WARREN W C.
XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-110599/10.
XX
PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.
XX
PS Claim 2; SEQ ID NO 8587; 245pp; English.
XX
CC The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridise to a second
CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non- translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 15112 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridisation between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and (b)
CC detecting the level or pattern of the complementary nucleic acid and the
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMFD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
SQ Sequence 464 BP; 132 A; 85 C; 107 G; 140 T; 0 U; 0 Other;

Query Match 1.9%; Score 25; DB 7; Length 464;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1302 GATGTCAAAAAA 1326

Db 42 GATGTCACAAAAA...AAAAA 18

RESULT 45
AAQ74392
ID AAQ74392 standard; DNA; 469 BP.

XX AC AAQ74392;

XX DT 25-MAR-2003 (revised)

XX DT 03-JUN-1995 (first entry)

XX DE Clone pSP65-Xal-4 encoding leech factor Xa inhibitor.

XX KW Leech; factor Xa; inhibitor; ss.

XX OS Hirudo medicinalis.

XX PH Key Location/Qualifiers
XX FT CDS 10..469
XX FT /*tag= a

XX PN W09423735-A1.

XX PD 27-OCT-1994.

XX XX 08-APR-1994; 94WO-US003871.

XX XX 09-APR-1993; 93US-00045805.

XX XX (BIOT-) BIO-TECHNOLOGY GENERAL CORP.

XX XX Zeelon EP, Werber MM, Levanon A;

XX XX WPI; 1994-341477/42.

XX XX P-PSDB; AAR62516.

XX PT Polypeptide from Hirudo medicinalis - has factor Xa inhibitory activity,
XX PT useful for treating blood coagulation disorders.

XX PS Example; Fig 5; 135pp; English.

XX CC Total RNA was extracted from 120 leeches, Hirudo medicinalis. An aliquot
XX CC of the poly A+ mRNA was used as template in a reverse transcription in
XX CC the presence of the synthetic primer A (AAQ74385) which provides
XX CC complementarity to the polyA sequence of the various mRNAs. Following
XX CC synthesis of the ss cDNA, the mRNA was degraded. An aliquot of the
XX CC neutralised ss cDNA was then subjected to PCR amplification using as
XX CC reverse primer the synthetic degenerative DNA oligomer B (AAQ74394). This
XX CC synthetic primer was prepd. in accordance with the first nine N-terminal
XX CC AAs of the naturally occurring Fxa inhibitor and encodes the first nine N
XX CC -terminal AAs in AAR62518. The PCR amplification products were loaded
XX CC onto agarose gel. Three distinct bands of about 350bp, 450 bp and 700 bp
XX CC were observed. The bands were hybridised to a synthetic radiolabeled DNA
XX CC probe (probe C) corresp. to N-terminal AAs 14-19 of AAR62518. The three
XX CC PCR products were found to hybridise with probe C. However, the band
XX CC corresp. to 700 bp was found to hybridise relatively poorly to the 350 bp
XX CC and 450 bp fragments. The DNA was purified, digested, and subcloned into
XX CC pSP65 which was used to transform E.coli strain MC1061. Transformants
XX CC were screening using probe C. Plasmid DNA was prepd. from positive
XX CC clones. Two classes of clones were obtd.: (a) those with an insert of
XX CC about 290 bps; and (b) those with an insert of about 450 bps. pSP65-Xal-4
XX CC contains an insert of about 450 bps. Its DNA sequence and deduced AA
XX CC sequence are given in AAQ74392/R62516. (Updated on 25-MAR-2003 to correct
XX CC PN field.)

XX Sequence 469 BP; 187 A; 70 C; 91 G; 121 T; 0 U; 0 Other;

Query Match 1.9%; Score 25; DB 2; Length 469;

Best Local Similarity 100.0%; Pred. No. 7.3;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA...AAAAA 1327
Db 445 ATGTCACAAAAA...AAAAA 469

RESULT 46

AAQ72957

ID AAQ72957 standard; DNA; 469 BP.

XX AC AAQ72957;

XX DT 27-AUG-2003 (revised)

XX DT 25-MAR-2003 (revised)

XX DT 04-JUN-1995 (first entry)

XX DE Leech PCR clone pSP65-Xal-4 encoding factor Xa inhibitor (FxaI).

XX KW Leech; factor Xa inhibitor; ss.

XX OS Hirudo medicinalis.

XX PH Key Location/Qualifiers
XX FT CDS 10..468
XX FT /*tag= a

XX PN W09423709-A1.

XX PD 27-OCT-1994.

XX XX 08-APR-1994; 94WO-US003918.

XX XX 09-APR-1993; 93US-00045804.

XX XX (BIOT-) BIO-TECHNOLOGY GENERAL CORP.
XX XX (YISS) YISSUM RES & DEV CO.

XX XX Werber MM, Zeelon EP, Levanon A, Guy R, Goldlust A, Rigbi M;
XX XX Panet A, Fischer M;

XX XX WPI; 1994-341457/42.

XX XX P-PSDB; AAR62624.

XX PT Recombinant factor 10a inhibitor of Hirudo medicinalis - for treating
XX PT excessive blood coagulation, partic. thrombosis, also related DNA,
XX PT vectors, transformed cells and antibodies.

XX PS Example; Fig 8; 107pp; English.

XX CC Total RNA was extracted from 120 leeches, poly A+ mRNA was isolated and
XX CC used as template in a reverse transcription reaction in the presence of
XX CC primer A (AAQ72958), which provides complementarity to the poly A
XX CC sequence of the various mRNAs. Following synthesis of the ss cDNA, the
XX CC mRNA was degraded and the ss cDNA subjected to PCR amplification using as
XX CC reverse primer the degenerative oligo B (AAQ72959) which was prepd. in
XX CC accordance with the first nine N-terminal AAs of factor Xa inhibitor
XX CC (FxaI) extracted from leech saliva. The PCR amplification products were
XX CC loaded onto agarose gel, the bands blotted onto nitrocellulose paper and
XX CC hybridised to a synthetic radiolabeled DNA probe (probe C (AAQ72962))
XX CC corresp. to N-terminal AAs 14-19 of leech saliva FxaI. Probe C was also
XX CC used to screen transformants of the PCR amplified DNA. The DNA sequence
XX CC and deduced AA sequence of one of the positive clones pSP65-Xal-4 (clone
XX CC 4) is shown in AAQ72957/R62624. It is noted that met 5 in AAR62624 is the
XX CC fourth AA (Met 1 is added by the bacterial host cell) of the FxaI
XX CC polypeptide which differs from Ile4 which is the fourth AA of the N-
XX CC terminal sequence of the naturally occurring FxaI isolated from leech
XX CC saliva. This discrepancy is apparently due to a mistake by the DNA
XX CC polymerase in the course of the PCR. (Updated on 25-MAR-2003 to correct
XX CC PN field.) (Updated on 27-AUG-2003 to correct OS field.)

XX Sequence 469 BP; 187 A; 70 C; 91 G; 121 T; 0 U; 0 Other;

Query Match 1.9%; Score 25; DB 2; Length 469;

Best Local Similarity 100.0%; Pred. No. 7.3;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
 |||||
 Db 445 ATGTCACAAAAA 469

RESULT 47
 AAV33944
 ID AAV33944 standard; DNA; 469 BP.
 XX
 AC AAV33944;
 XX
 DT 25-JAN-1999 (first entry)
 XX
 DE Leech FXaI gene clone pSP65-XaI-4.
 XX
 XX European leech; Factor Xa inhibitor; FXaI; probe; hybridisation; RT-PCR;
 KW influenza virus; infection; coagulation; primer; amplification; ss.
 XX
 OS Hirudo medicinalis.
 XX

Key Location/Qualifiers
 CDS 10..468
 FT /*tag= a
 FT /transl_except= (pos:268..270, aa:Xaa)
 FT /transl_except= (pos:280..282, aa:Xaa)
 FT /transl_except= (pos:328..330, aa:Xaa)
 FT /transl_except= (pos:334..336, aa:Xaa)
 FT /transl_except= (pos:391..393, aa:Xaa)
 FT /note= "Xaa is unknown and not given in the corresponding
 protein sequence"
 XX
 PN US5824641-A.
 XX
 PD 20-OCT-1998.
 XX
 XX 06-JUN-1995; 95US-00467389.
 XX
 PR 09-APR-1993; 93US-00045805.
 PR 08-APR-1994; 94US-00225442.
 XX
 PA (BIOT-) BIO-TECHNOLOGY GENERAL CORP.
 XX
 XX Werber MM;
 PI
 DR WPI; 1998-582592/49.
 DR P-PSDB; AAW68544.
 XX
 PT Treating or preventing influenza - comprises the administration of factor
 PT Xa inhibiting polypeptide.
 XX
 PS Example 3; Fig 5; 56pp; English.
 XX
 CC This sequence represents clone pSP65-XaI-4 which was amplified from DNA
 CC isolated from the European leech (Hirudo medicinalis). The open reading
 CC frame contains 5 internal stop codons when decoded from the first ATG
 CC initiation codon. No equivalent residues are given in the corresponding
 CC protein sequence. The fragment is a primary isolated and partially
 CC encoded the leech Factor Xa inhibitor (FXaI). The fragment was amplified
 CC by the primers AAV33941-V33942. The fragment was used as a probe to
 CC screen a cDNA library for the FXaI gene (AAV33940). The FXaI protein is
 CC claimed to be useful in the treatment of influenza viral infection by
 CC preventing the infection and re-infection cycle of cells by the virus. An
 CC activation enzyme implicated in the viral infection cycle has been shown
 CC to be similar to the chicken coagulation Factor Xa and it is thought that
 CC the corresponding human protein may act in a similar manner. Thus
 CC inhibitors of Factor Xa may prevent influenza infections
 XX
 SQ Sequence 469 BP; 187 A; 70 C; 91 G; 121 T; 0 U; 0 Other;

Query Match 1.9%; Score 25; DB 2; Length 469;
 Best Local Similarity 100.0%; Pred. No. 7.3;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
 |||||
 Db 445 ATGTCACAAAAA 469

RESULT 48
 AAX01821
 ID AAX01821 standard; DNA; 469 BP.
 XX
 AC AAX01821;
 XX
 DT 13-APR-1999 (first entry)
 XX
 DE H. medicinalis FXaI clone pSP65-XaI-4 DNA.
 XX
 KW Factor Xa; FXaI; inhibitor; blood; coagulation; European leech;
 KW post-operative trauma; obesity; pregnancy; oral contraceptive; stroke;
 KW cerebrovascular disorder; ss.
 XX
 OS Hirudo medicinalis.
 XX

Key Location/Qualifiers
 CDS 10..469
 FT /*tag= a
 FT /product= "pSP65-XaI-4"
 FT /note= "contains internal stop codons"
 XX
 PN US5863534-A.
 XX
 PD 26-JAN-1999.
 XX
 XX 06-JUN-1995; 95US-00469219.
 XX
 PR 09-APR-1993; 93US-00045805.
 PR 08-APR-1994; 94US-00225442.
 XX
 PA (BIOT-) BIO-TECHNOLOGY GENERAL CORP.
 XX
 XX Werber MM, Levanon A, Zeelon EP;
 PI
 DR WPI; 1999-131254/11.
 DR P-PSDB; AAW92370.
 XX
 PT Reducing blood coagulation - using factor Xa inhibitor polypeptides
 PT obtained from saliva of the European leech Hirudo medicinalis.
 XX
 PS Disclosure; Fig 5; 56pp; English.
 XX
 CC This sequence encodes a novel FXaI protein, isolated from the European
 CC leech (Hirudo medicinalis) clone pSP65-XaI-4 which inhibits Factor Xa
 CC activity. This fragment is used in the construction of a polypeptide
 CC which can be used to reduce the extent of blood coagulation and has the
 CC sequence X-V-Cys Gln Glu Glu Cys Pro Asp Pro Tyr Leu Cys Ser Pro Val
 CC Thr Asn Arg Cys Glu Cys Thr Pro Val Leu Cys Arg Met Tyr Cys Lys Phe Trp
 CC Ala Lys Cys Glu Glu Cys Gln Ser Val Thr Asn Arg Cys Asp Cys Gln
 CC Asp Phe Lys Cys Pro Gln Ser Tyr Cys-Z, where X is Met or absent; Y is 0-
 CC 29 amino acids of sequence Lys Met Cys Trp Asn Lys Gly Cys Pro Cys Gly
 CC Gln Arg Cys Asn Leu His Arg Asn Glu Cys Glu Val Ile Ala Glu Asn Ile Glu
 CC with the proviso that if Y is > 0 amino acids, then the carboxy-terminal
 CC sequence of Y beginning with Glu as the carboxy-terminal amino acid is
 CC present and where Val24 may be preceded by Gly, Z is absent or all or a
 CC part of the sequence Pro110-Lys156 (Pro His Gly Phe Glu Thr Asp Glu Asn
 CC Thr Lys Lys Lys Pro Arg Thr Ile Asp Arg Leu Lys Asn Trp Phe Lys Lys
 CC Phe Gly Lys), with the proviso that if Z is > 0 amino acid, then the
 CC amino-terminal sequence of Z beginning with Pro as the amino-terminal
 CC amino acid is present. The method can be used for treating subjects with
 CC excessive blood coagulation disorders, e.g. vascular disorders, post-
 CC operative trauma, obesity, pregnancy, side effects of oral
 CC contraceptives, prolonged immobilisation, or a cerebrovascular disorder
 CC e.g. stroke

```
XX SQ Sequence 469 BP; 187 A; 70 C; 91 G; 121 T; 0 U; 0 Other;
Query Match 1.9%; Score 25; DB 2; Length 469;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 445 ATGTCACAAAAA 469

RESULT 49
AAV55665
ID AAV55665 standard; DNA; 469 BP.
XX AC AAV55665;
XX DT 18-MAR-1999 (first entry)
XX DE Factor Xa inhibitor peptide coding sequence.
XX KW Factor Xa inhibitor; leech; blood coagulation; thrombosis;
XX KW recurrent influenza infection; ss.
XX OS Hirudo medicinalis.
XX PN US5858970-A.
XX PD 12-JAN-1999.
XX PF 07-JAN-1997; 97US-00779379.
XX PR 09-APR-1993; 93US-00045805.
XX PR 08-APR-1994; 94US-00225442.
XX PA (BIOT-) BIO-TECHNOLOGY GENERAL CORP.
XX PI Levanon A, Zeelon EP, Werber MM;
XX WPI; 1999-119954/10.
XX P-PSDB; AAW73598.
XX Factor Xa inhibitory polypeptide - derived from Hirudo medicinalis,
XX useful as anticoagulant or antithrombotic agent.
XX Example 3; Fig 5; 57pp; English.
XX CC This sequence encodes the full length factor Xa inhibitor peptide of the
XX invention. The factor Xa sequence was isolated from the leech Hirudo
XX medicinalis. It can be used to reduce the extent of blood coagulation or
XX thrombosis or to prevent recurrent influenza infection
XX SQ Sequence 469 BP; 187 A; 70 C; 91 G; 121 T; 0 U; 0 Other;
Query Match 1.9%; Score 25; DB 2; Length 469;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 445 ATGTCACAAAAA 469

RESULT 50
AAD04138
ID AAD04138 standard; DNA; 469 BP.
XX AC AAD04138;
XX DT 02-JUL-2001 (first entry)
XX DE DNA clone pSP65-Xai-4 encoding leech FXa inhibitor.
XX SQ Sequence 469 BP; 187 A; 70 C; 91 G; 121 T; 0 U; 0 Other;

XX KW FXaI; blood coagulation; factor FXa; inhibitor; therapy; thrombosis;
XX influenza; leech; ds.
XX Hirudo medicinalis.
XX Key Location/Qualifiers
XX CDS 10..453
XX /tag= a
XX /product= "Leech FXa inhibitor encoded by clone pSP65-Xai-4"
XX /transl_except= (pos:265..273, aa:Gly-Thr)
XX /note= "Insertion of an in-frame stop codon alters the reading frame"
XX /transl_except= (pos:277..285, aa:Ile-Lys)
XX /note= "Insertion of an in-frame stop codon alters the reading frame"
XX /transl_except= (pos:325..333, aa:Leu-Arg)
XX /note= "Insertion of an in-frame stop codon alters the reading frame"
XX /transl_except= (pos:331..339, aa:Arg-Ile)
XX /note= "Insertion of an in-frame stop codon alters the reading frame"
XX /transl_except= (pos:388..396, aa:Ser-Ile)
XX /note= "Insertion of an in-frame stop codon alters the reading frame; All the above translation exceptions correspond to the sequence shown in AAE00819"
XX /transl_except= (pos:268..270, aa:Xaa)
XX /transl_except= (pos:280..282, aa:Xaa)
XX /transl_except= (pos:328..330, aa:Xaa)
XX /transl_except= (pos:334..336, aa:Xaa)
XX /note= "Xaa corresponds to in-frame stop codon; the above translation exceptions correspond to the sequence shown in AAE00755; CDS does not include stop codon"
XX /partial

XX PN US6211341-B1.
XX PD 03-APR-2001.
XX PF 11-JAN-1999; 99US-00228152.
XX PR 09-APR-1993; 93US-00045805.
XX PR 08-APR-1994; 94US-00225442.
XX PR 06-JUN-1995; 95US-00469219.
XX PA (BIOT-) BIO-TECHNOLOGY GEN CORP.
XX PI Zeelon EP, Werber MM, Levanon A;
XX WPI; 2001-280869/29.
XX P-PSDB; AAE00755, AAE00818.
XX New antibody reactive with an epitope of a polypeptide having factor Xa inhibitory activity, useful for purifying or detecting a polypeptide having Factor Xa inhibitory activity.
XX Example 3; Col 51-54; 57pp; English.
XX The invention relates to an antibody which specifically reacts with an epitope of a novel (unknown) polypeptide derived from Hirudo medicinalis. This polypeptide is an inhibitor of blood coagulation factor FXa and is designated as FXaI. The polypeptide of the invention is used in diagnostic and therapeutic methods applied in conditions characterised by the excessive blood coagulation and thrombosis. The FXaI polypeptide is also used to prevent recurrent influenza infection. The monoclonal antibody specific for FXaI polypeptide is useful for purifying or detecting FXaI. The present sequence is DNA clone pSP65-Xai-4 encoding leech FXa inhibitor
XX SQ Sequence 469 BP; 187 A; 70 C; 91 G; 121 T; 0 U; 0 Other;
```


Query Match 1.9%; Score 25; DB 4; Length 469;
 Best Local Similarity 100.0%; Pred. No. 7.3;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCAAAAAAAAAAAAAAAAA 1327
 ||||||||||||||||||
 Db 445 ATGTCAAAAAAAAAAAAAAAAA 469

Search completed: September 17, 2004, 03:18:25
 Job time : 564 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 03:09:09 ; Search time 3727 Seconds
(without alignments)
1797.146 Million cell updates/sec

Title: US-09-927-091-3_COPY_2500_3826

Perfect score: 1327

Sequence: 1 tccacagtgtcacagtag.....aaaaaaaaaaaaaaaaaaaa 1327

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 3327077 seqs, 2523723180 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6654154

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCPUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09D_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09E_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1327	100.0	3826	9	US-09-927-091-3
2	1308	98.6	23433	9	US-09-927-091-7
3	1257	94.7	30676	9	US-09-927-091-8
4	726	54.7	30625	9	US-09-927-091-5
5	420	31.7	610	13	US-10-027-632-100255
6	420	31.7	610	16	US-10-027-632-100265
7	60	4.5	60	10	US-09-908-975-16197
8	29	2.2	1652	16	US-10-116-275-299
9	27	2.0	1754	17	US-10-437-963-80513
10	27	2.0	248436	13	US-10-087-192-2014
11	26	2.0	273	12	US-09-969-034-4195
12	26	2.0	393	9	US-09-960-352-6255
13	26	2.0	411	9	US-09-960-352-13315
14	26	2.0	4702	16	US-10-295-027-141
15	26	2.0	4702	16	US-10-295-027-141

15	26	2.0	4702	16	US-10-295-027-1164	Sequence 1164, Ap
16	26	2.0	6808	17	US-10-437-963-87284	Sequence 87284, A
17	25	1.9	136	13	US-10-085-783A-47925	Sequence 47925, A
18	25	1.9	136	16	US-10-242-535A-47925	Sequence 47925, A
19	25	1.9	143	13	US-10-085-783A-51657	Sequence 51657, A
20	25	1.9	143	16	US-10-242-535A-51657	Sequence 51657, A
21	25	1.9	156	13	US-10-085-783A-45661	Sequence 45661, A
22	25	1.9	156	16	US-10-242-535A-45661	Sequence 45661, A
23	25	1.9	184	9	US-09-867-701-9253	Sequence 9253, Ap
24	25	1.9	191	9	US-09-962-436-134	Sequence 134, App
25	25	1.9	191	9	US-09-954-456-1020	Sequence 1020, Ap
26	25	1.9	191	9	US-09-880-107-525	Sequence 525, App
27	25	1.9	207	9	US-08-867-701-10779	Sequence 10779, A
28	25	1.9	222	9	US-09-880-701-587	Sequence 587, App
29	25	1.9	256	13	US-10-424-599-31507	Sequence 31507, A
30	25	1.9	296	13	US-10-085-783A-52889	Sequence 52889, A
31	25	1.9	296	16	US-10-242-535A-52889	Sequence 52889, A
32	25	1.9	330	9	US-09-813-358-64	Sequence 64, Appl
33	25	1.9	330	10	US-09-997-279-64	Sequence 64, Appl
34	25	1.9	354	12	US-09-922-293-42	Sequence 42, Appl
35	25	1.9	354	9	US-09-960-352-1456	Sequence 1456, Ap
36	25	1.9	358	13	US-10-424-599-52765	Sequence 52765, A
37	25	1.9	382	17	US-10-437-963-88420	Sequence 88420, A
38	25	1.9	390	13	US-10-175-184A-31	Sequence 31, Appl
39	25	1.9	430	10	US-09-918-995-7699	Sequence 7699, Ap
40	25	1.9	464	9	US-09-960-352-8587	Sequence 8587, Ap
41	25	1.9	467	15	US-10-198-846-1221	Sequence 1221, Ap
42	25	1.9	483	15	US-10-102-524-36	Sequence 36, Appl
43	25	1.9	483	17	US-10-437-963-85418	Sequence 85418, A
44	25	1.9	486	13	US-10-027-632-93705	Sequence 93705, A
45	25	1.9	486	13	US-10-027-632-305902	Sequence 305902, A
46	25	1.9	486	16	US-10-027-632-95705	Sequence 95705, A
47	25	1.9	486	16	US-10-027-632-305902	Sequence 305902, A
48	25	1.9	492	17	US-10-437-963-19508	Sequence 19508, A
49	25	1.9	503	15	US-10-029-386-1201	Sequence 1201, Ap
50	25	1.9	527	9	US-09-925-301-787	Sequence 787, App
51	25	1.9	527	12	US-09-969-034-4262	Sequence 4262, Ap
52	25	1.9	558	17	US-10-437-963-79631	Sequence 79631, A
53	25	1.9	591	10	US-09-764-891-1170	Sequence 1170, Ap
54	25	1.9	593	13	US-10-027-632-96478	Sequence 96478, A
55	25	1.9	593	16	US-10-027-632-96478	Sequence 96478, A
56	25	1.9	613	17	US-10-437-963-17992	Sequence 17992, A
57	25	1.9	616	13	US-10-027-632-308292	Sequence 308292, A
58	25	1.9	616	13	US-10-027-632-308293	Sequence 308293, A
59	25	1.9	616	13	US-10-027-632-308294	Sequence 308294, A
60	25	1.9	616	13	US-10-027-632-308295	Sequence 308295, A
61	25	1.9	616	16	US-10-027-632-308292	Sequence 308292, A
62	25	1.9	616	16	US-10-027-632-308293	Sequence 308293, A
63	25	1.9	616	16	US-10-027-632-308294	Sequence 308294, A
64	25	1.9	616	16	US-10-027-632-308295	Sequence 308295, A
65	25	1.9	617	15	US-10-198-846-8111	Sequence 8111, Ap
66	25	1.9	626	13	US-10-027-632-192324	Sequence 192324, A
67	25	1.9	626	16	US-10-027-632-192324	Sequence 192324, A
68	25	1.9	645	13	US-09-823-245A-41	Sequence 41, Appl
69	25	1.9	648	15	US-10-198-846-8765	Sequence 8765, Ap
70	25	1.9	651	17	US-10-437-963-92674	Sequence 92674, A
71	25	1.9	708	17	US-10-767-701-11186	Sequence 11186, A
72	25	1.9	730	13	US-10-027-632-100233	Sequence 100233, A
73	25	1.9	730	16	US-10-027-632-100233	Sequence 100233, A
74	25	1.9	732	17	US-10-437-963-46535	Sequence 46535, A
75	25	1.9	749	13	US-10-027-632-20596	Sequence 20596, A
76	25	1.9	749	16	US-10-027-632-20596	Sequence 20596, A
77	25	1.9	834	13	US-10-027-632-165337	Sequence 165337, A
78	25	1.9	834	13	US-10-027-632-165338	Sequence 165338, A
79	25	1.9	834	16	US-10-027-632-165337	Sequence 165337, A
80	25	1.9	834	16	US-10-027-632-165338	Sequence 165338, A
81	25	1.9	1026	17	US-10-437-963-50383	Sequence 50383, A
82	25	1.9	1092	9	US-09-925-301-317	Sequence 317, App
83	25	1.9	1163	17	US-10-437-963-9930	Sequence 9930, Ap
84	25	1.9	1193	13	US-10-424-599-75445	Sequence 75445, A
85	25	1.9	1307	15	US-10-198-846-10951	Sequence 10951, A
86	25	1.9	1410	17	US-10-437-963-10738	Sequence 10738, A
87	25	1.9	1451	15	US-10-106-698-401	Sequence 401, App

QY 1081 TCTCTGACAAACACACACATCCCAAGTAGTCCCGAAGAGCTAAACACAGGGGTTCTTAA 1140
Db 14630 TCTCTGACAAACACACACATCCCAAGTAGTCCCGAAGAGCTAAACACAGGGGTTCTTAA 14689
QY 1141 AATGGCTGCCCGCCGACCCGGGCTCCCTTTGGSCAAAAGGAATGTCAGCCCTACCCCA 1200
Db 14690 AATGGCTGCCCGCCGACCCGGGCTCCCTTTGGSCAAAAGGAATGTCAGCCCTACCCCA 14749
QY 1201 ACCCTTCAACTACCAGAAATCTGGGCCACCCAGCAGTATTTTATTTAAATGTTGCCCA 1260
Db 14750 ACCCTTCAACTACCAGAAATCTGGGCCACCCAGCAGTATTTTATTTAAATGTTGCCCA 14809
QY 1261 TTTTATGAGTTATGATCAATTTGTTATTAATTTAAAGTTACAGATGTCA 1308
Db 14810 TTTTATGAGTTATGATCAATTTGTTATTAATTTAAAGTTACAGATGTCA 14857

RESULT 3

US-09-927-091-8
; Sequence 8, Application US/09927091
; Patent No. US20020119541A1
; GENERAL INFORMATION:
; APPLICANT: KILLARY, ANN
; APPLICANT: LOTT, STEVE
; APPLICANT: CHANDLER, DAWN
; TITLE OF INVENTION: THE TUMOR SUPPRESSOR CAR-1
; FILE REFERENCE: UTSC:651US
; CURRENT APPLICATION NUMBER: US/09/927,091
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/227,560
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 60/225,033
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 30676
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (6671)..(30676)
; OTHER INFORMATION: n = A or C or G or T/U
US-09-927-091-8

Query Match 94.7%; Score 1257; DB 9; Length 30676;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1307; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TCCACAGTGGTCCACAGGTAGTACCTGGTCTAGGTTGGCTGAGAGCCAACTCTCCTGC 60
Db 24440 TCCACAGTGGTCCACAGGTAGTACCTGGTCTAGGTTGGCTGAGAGCCAACTCTCCTGC 24499
QY 61 CACCCCCACACCAAGAACTATATGTTTCTACTCTCCACATGATCTGTGTCAGTAT 120
Db 24500 CACCCCCACACCAAGAACTATATGTTTCTACTCTCCACATGATCTGTGTCAGTAT 24559
QY 121 GATGCTGTGGCTGTGGAGGCACTGGTGTAGTCCACATATATAGTCATGTGCA 180
Db 24560 GATGCTGTGGCTGTGGAGGCACTGGTGTAGTCCACATATATAGTCATGTGCA 24619
QY 181 CCACCTTCCTGCCACAGCCGAGGACAGGTGTGGGTATACCCAAAGCTGATCAGAG 240
Db 24620 CCACCTTCCTGCCACAGCCGAGGACAGGTGTGGGTATACCCAAAGCTGATCAGAG 24679
QY 241 CCCATTAGCTTAAAGCAACTCCAGACAAAGCCTCCCTGGATGATCGAGTCCCGCAGTAG 300
Db 24680 CCCATTAGCTTAAAGCAACTCCAGACAAAGCCTCCCTGGATGATCGAGTCCCGCAGTAG 24739
QY 301 CTCTGAACAAGAGTCCAGCCAAACCTCTTTCAGCCAGGCTCTGTGACCTGTAGGGTGCA 360
Db 24740 CTCTGAACAAGAGTCCAGCCAAACCTCTTTCAGCCAGGCTCTGTGACCTGTAGGGTGCA 24799

QY 361 GGAGGCTTCCAGAGCAGTTGTTGTAATTAGGACCCCAAGCACCTGGAGGGGCTGTTGGCT 420
Db 24800 GGAGGCTTCCAGAGCAGTTGTTGTAATTAGGACCCCAAGCACCTGGAGGGGCTGTTGGCT 24859
QY 421 AGACCCCTTGTCCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTCAGAAAAACAAGAC 480
Db 24860 GGACCCCTTGTCCAGACTTGGCATCTATCTCAGTTAGGATCCTGCTCAGAAAAACAAGAC 24919
QY 481 CACTGTGAGCTGGTTTAAATTTAGACAAGGATTTACTACCTGGCCCTGGTGGCTTGC AAAA 540
Db 24920 CACTGTGAGCTGGTTTAAATTTAGACAAGGATTTACTACCTGGCCCTGGTGGCTTGC AAAA 24979
QY 541 TTGTTGAACAGACTGGAGAAGACAGACTCTCTGAAATTTCCAGGAATCTCCAGCCGCCAGAT 600
Db 24980 TTGTTGAACAGACTGGAGAAGACAGACTCTCTGTAATTTCCAGGAATCTCCAGCCGCCAGAT 25039
QY 601 TCATCATGTCTGTTGTGACCAAGAAAGCTGCCCATCTCTGAGGAAGCCACTATGCCAGA 660
Db 25040 TCATCATGTCTGTTGTGACCAAGAAAGCTGCCCATCTCTGAGGAAGCCACTATGCCAGA 25099
QY 661 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGCTCCGTGCCAGCCCAATAGATGTC 720
Db 25100 AAGCTGCTGACTGCAGAACTAGGCTCCCTCTGCCACGCTCCGTGCCAGCCCAATAGATGTC 25159
QY 721 CTGAGGCTGCCCTCTCCCACTTCACTCAGTTTCCCAATCTAAATTTTACAAGAGATT 780
Db 25160 CTGAGGCTGCCCTCTCCCACTTCACTCAGTTTCCCAATCTAAATTTTACAAGAGATT 25219
QY 781 CTGTTTGGGGAACTTAAGTCAGATCCAGAACTTGGCTGCAAGGGAGTCTGGGAAATGT 840
Db 25220 CTGTTTGGGGAACTTAAGTCAGATCCAGAACTTGGCTGCAAGGGAGTCTGGGAAATGT 25279
QY 841 CATTTCCCTAGAGGAAGTGTAGGCTGGGTGAGCAAGCCACCTGCTGCTTTTCTGCCAC 900
Db 25280 CATTTCCCTAGAGGAAGTGTAGGCTGGGTGAGCAAGCCACCTGCTGCTTTTCTGCCAC 25339
QY 901 AGCATCCAATCGTGAAGAACTCGGAGAGGGTGGAGTCCACATCTAGGGTTGTCTGCC 960
Db 25340 AGCATCCAATCGTGAAGAACTCGGAGAGGGTGGAGTCCACATCTAGGGTTGTCTGCC 25399
QY 961 CTGCTCTATATCCTGCCAGAGGTGGAACTGAGAGAGTGGGTGCAAGACTGAGCCTA 1020
Db 25400 CTGCTCTATATCCTGCCAGAGGTGGAACTGAGAGAGTGGGTGCAAGACTGAGCCTA 25459
QY 1021 AATGTCCTCCCGGCTTGCATTTCTTCTAGTCTCGGGCTAGATTCTGCACCTGGG 1080
Db 25460 AATGTCCTCCCGGCTTGCATTTCTTCTAGTCTCGGGCTAGATTCTGCACCTGGG 25519
QY 1081 TCTCTGACACACACACCATCCCAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 1140
Db 25520 TCTCTGACACACACACCATCCCAAGTAGCCGGAAGAGCTAAACACAGGGGGTTCTTAA 25579
QY 1141 AATGCTGCCCGGCCACCCGGGCTCCCTTTGGGCAAAAGAAATGTGACGCCCTACCCCA 1200
Db 25580 AATGCTGCCCGGCCACCCGGGCTCCCTTTGGGCAAAAGAAATGTGACGCCCTACCCCA 25639
QY 1201 ACCCTTCAACTACAGAACTGGGCCACCCAGCAGTATTTTATTTAAATGTTGCCCA 1260
Db 25640 ACCCTTCAACTACAGAACTGGGCCACCCAGCAGTATTTTATTTAAATGTTGCCCA 25699
QY 1261 TTTTATGAGTTATGATCAATTTGTTATTAATTTAAAGTTACAGATGTCA 1308
Db 25700 TTTTATGAGTTATGATCAATTTGTTATTAATTTAAAGTTACAGATGTCA 25747

RESULT 4

US-09-927-091-5
; Sequence 5, Application US/09927091
; Patent No. US20020119541A1
; GENERAL INFORMATION:
; APPLICANT: KILLARY, ANN
; APPLICANT: LOTT, STEVE

```

; APPLICANT: CHANDLER, DAWN
; TITLE OF INVENTION: THE TUMOR SUPPRESSOR CAR-1.
; FILE REFERENCE: UTSC:651US
; CURRENT APPLICATION NUMBER: US/09/927,091
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/227,560
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 60/225,033
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 30625
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (4754)..(30625)
; OTHER INFORMATION: n = A or C or G or T/U
US-09-927-091-5

Query Match          54.7%; Score 726; DB 9; Length 30625;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1146; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 157 TCACACATTATAGTCATGTGCCACACACCTTCTGCCCCCAGAGCCGAGGACAGGGTGAG 216
Db 23131 TCACACATTATAGTCATGTGCCACACCTTCTGCCCCCAGAGCCGAGGACAGGGTGAG 23190

QY 217 GGTATACCCAAAGCTGATGACAGGCCCATTTAGGCTTAAAGCAACTGCGAGCAAGCTCC 276
Db 23191 GGTATACCCAAAGCTGATGACAGGCCCATTTAGGCTTAAAGCAACTGCGAGCAAGCTCC 23250

QY 277 CTGGATGATCGAGGTCCCGAGTGTCTGACAGAGCTGACAGCAACCTCTTCAGCCAG 336
Db 23251 CTGGATGATCGAGGTCCCGAGTGTCTGACAGAGCTGACAGCAACCTCTTCAGCCAG 23310

QY 337 GCCTCTGTGACCTGTAGGTGCGAGGCTTCCAGAGCAGTGTGTAAATTAGGACCC 396
Db 23311 GCCTCTGTGACCTGTAGGTGCGAGGCTTCCAGAGCAGTGTGTAAATTAGGACCC 23370

QY 397 AAGCACT-GGGAGGGCTGTGCTAGACCCCTGTGACACTTGGCACTTATCTCAGTTA 455
Db 23371 AAGCACTGGGAGGGCTGTGCTAGACCCCTGTGACACTTGGCACTTATCTCAGTTA 23430

QY 456 GGATCTGTGCGAGAAACAGAGCCACTGTGAGCTGGTTTAAATTAGCAAGATTTACT 515
Db 23431 GGATCTGTGCGAGAAACAGAGCCACTGTGAGCTGGTTTAAATTAGCAAGATTTACT 23490

QY 516 ACCTGCCCCCTGTGCTGCTTGCMAAATTTGTTGAAGAGCTGGAGAAGCAGACTCTGCTGAA 575
Db 23491 ACCTGCCCCCTGTGCTGCTTGCMAAATTTGTTGAAGAGCTGGAGAAGCAGACTCTGCTGAA 23550

QY 576 TTTCAGGAACCTCCAGGCGCAGATTCATCATCTGTTGTGTGACCGAGAAAGCTGCCCCC 635
Db 23551 TTTCAGGAACCTCCAGGCGCAGATTCATCATCTGTTGTGTGACCGAGAAAGCTGCCCCC 23610

QY 636 ATCTGCGAGAGCCACTATGCCAGAGCTGCTGACTGCAAGACTAGGCTCCCTCTGCGCA 695
Db 23611 ATCTGCGAGAGCCACTATGCCAGAGCTGCTGACTGCAAGACTAGGCTCCCTCTGCGCA 23670

QY 696 CGGTCCGTGCGAGCCCAATAGATCTCTGAGGCCCTGCCCCCTCTCCCACTTCACTCAGTTCC 755
Db 23671 CGGTCCGTGCGAGCCCAATAGATCTCTGAGGCCCTGCCCCCTCTCCCACTTCACTCAGTTCC 23730

QY 756 CAAATCTAAATTTTTCAGAGATTCTGTTGGGGAACTTAAGTCAGATCCAGAACCTT 815
Db 23731 CAAATCTAAATTTTTCAGAGATTCTGTTGGGGAACTTAAGTCAGATCCAGAACCTT 23790

QY 816 GCGTCAAGGGAGTCTGGGAAATGTCATTCCCTAGAGAAAGTTAGGTGGGTGGAGCA 875
Db 23791 GCGTCAAGGGAGTCTGGGAAATGTCATTCCCTAGAGAAAGTTAGGTGGGTGGAGCA 23850

```

```

QY 876 AGCCCCACCTGCGTTTTTCTGCCCACAGCATCCAACTCGTAGACTCGGGAGAGGGTGA 935
Db 23851 AGCCCCACCTGCGTTTTTCTGCCCACAGCATCCAACTCGTAGAAACTCGGGAGAGGGTGA 23910

QY 936 GTCCACATCTAGGTTGTCTCTGCCCCCTTGGCTCTATCCCTGCCACAGAGTGGGAACCTGA 995
Db 23911 GTCCACATCTAGGTTGTCTCTGCCCCCTTGGCTCTATCCCTGCCACAGAGTGGGAACCTGA 23970

QY 996 GGAGTGGGCTGCAAGACTGAGCTGAAGTCTCTCCCGGCTTGAATTTCTTTCTAGTCC 1055
Db 23971 GGAGTGGGCTGCAAAACTGAGCTTAAATGTCTCCCGGCTTGAATTTCTTTCTAGTCC 24030

QY 1056 TGGGGCTAGATCTGACATTTGGGCTCTGACACACACACCATCCCAAGTAGCCGA 1115
Db 24031 TGGGGCTTAAATTTCTGACCTTGGGCTCTGACACACACACCATCCCAAGTAGCCGA 24090

QY 1116 AGAGCTAAACACAGGGGGTTCTTAAATGTGCTGCCCGGCCACCCGGGCTCTCCCTTGGGC 1175
Db 24091 AAAGCTAAACACAGGGGGTTCTTAAATGTGCTGCCCGGCCACCCGGGCTCTCCCTTGGGC 24150

QY 1176 AAAAGGAATTTGTCAGCCCTTACCCCAACCCCTTAACTACCAAGATCTGGGCAACCCAGCA 1235
Db 24151 AAAAGGAATTTGTCAGCCCTTACCCCAACCCCTTCAACTACCAAGATCTGGGCAACCCAGCA 24210

QY 1236 GTATTTTATTTAAATTTGTTGCCATTTTATGATTTATGATTTGATTTAAATTTAA 1295
Db 24211 GTATTTTATTTAAATTTGTTGCCATTTTATGATTTATGATTTGATTTAAATTTAA 24270

QY 1296 GTTACAGATGTCA 1308
Db 24271 GTTACAGATGTCA 24283

RESULT 5
US-10-027-632-100265/c
; Sequence 100265, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 100265
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-100265

Query Match          31.7%; Score 420; DB 13; Length 610;
Best Local Similarity 100.0%; Pred. No. 1.7e-02;
Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCACAGTGGTTCACAGGTAGTAGTACCTGGTCTTGGGTTGCTAGGGTTGCTGAGAGCCAACTCTCTCTGC 60
Db 455 TCACAGTGGTTCACAGGTAGTAGTACCTGGTCTTGGGTTGCTGAGAGCCAACTCTCTCTGC 396

```

QY 61 CACCCCAACACAGAGTATATAGTTCCTACTTCTCCCACTGATCTGCTGCTCAGTGAT 120
 Db 395 CACCCCAACACAGAGTATATAGTTCCTACTTCTCCCACTGATCTGCTGCTCAGTGAT 336
 QY 121 GATGCTGTGGCTGTGGAAGGCACTGGTGTAGTTCAGTCCACACATTATAGTTCATGTGCCA 180
 Db 335 GATGCTGTGGCTGTGGAAGGCACTGGTGTAGTTCAGTCCACACATTATAGTTCATGTGCCA 276
 QY 181 CCACCTTCTGTCACAGGCGGAGGACAGGGTGAGGTTATACCCAAAGCTGATGCAGAG 240
 Db 275 CCACCTTCTGTCACAGGCGGAGGACAGGGTGAGGTTATACCCAAAGCTGATGCAGAG 216
 QY 241 CCCATTAGCTTAAAGCAACTGCAGGACAGCTCCCTGGATGATCGAGGTCCTCCCACTAG 300
 Db 215 CCCATTAGCTTAAAGCAACTGCAGGACAGCTCCCTGGATGATCGAGGTCCTCCCACTAG 156
 QY 301 CTCTGAACAGAGTCCAGCCCAACCTCTTCAGCCAGGCTCTGTGACCTGCTAGGGTGCA 360
 Db 155 CTCTGAACAGAGTCCAGCCCAACCTCTTCAGCCAGGCTCTGTGACCTGCTAGGGTGCA 96
 QY 361 GGAGGCTTCCAGAGAGCTGTTGTAATTAGGACCCCAAGCACTGGGAGGGGCTGTGGCT 420
 Db 95 GGAGGCTTCCAGAGAGCTGTTGTAATTAGGACCCCAAGCACTGGGAGGGGCTGTGGCT 36

RESULT 6
 US-10-027-632-100265/c
 ; Sequence 100265, Application US/10027632
 ; Publication No. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE OF INVENTION: Polymorphisms in the Human Genome
 ; FILE REFERENCE: 108827.123
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; PRIOR FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 100265
 ; LENGTH: 610
 ; TYPE: DNA
 ; ORGANISM: Human
 US-10-027-632-100265

Query Match 31.7%; Score 420; DB 16; Length 610;
 Best Local Similarity 100.0%; Pred. No. 1.7e-202;
 Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TCCACAGTGGTTCACAGGTAGTACCTGGTCTAGGTTGCTGAGGACCAACCTCTCCCTGC 60
 Db 455 TCCACAGTGGTTCACAGGTAGTACCTGGTCTAGGTTGCTGAGGACCAACCTCTCCCTGC 396
 QY 61 CACCCCAACACAGAGTATATAGTTCCTACTTCTCCCACTGATCTGCTGCTCAGTGAT 120
 Db 395 CACCCCAACACAGAGTATATAGTTCCTACTTCTCCCACTGATCTGCTGCTCAGTGAT 336
 QY 121 GATGCTGTGGCTGTGGAAGGCACTGGTGTAGTTCAGTCCACACATTATAGTTCATGTGCCA 180

Db 335 GATGCTGTGGCTGTGGAAGGCACTGGTGTAGTTCAGTCCACACATTATAGTTCATGTGCCA 276
 QY 181 CCACCTTCTGTCACAGGCGGAGGACAGGGTGAGGTTATACCCAAAGCTGATGCAGAG 240
 Db 275 CCACCTTCTGTCACAGGCGGAGGACAGGGTGAGGTTATACCCAAAGCTGATGCAGAG 216
 QY 241 CCCATTAGCTTAAAGCAACTGCAGGACAGCTCCCTGGATGATCGAGGTCCTCCCACTAG 300
 Db 215 CCCATTAGCTTAAAGCAACTGCAGGACAGCTCCCTGGATGATCGAGGTCCTCCCACTAG 156
 QY 301 CTCTGAACAGAGTCCAGCCCAACCTCTTCAGCCAGGCTCTGTGACCTGCTAGGGTGCA 360
 Db 155 CTCTGAACAGAGTCCAGCCCAACCTCTTCAGCCAGGCTCTGTGACCTGCTAGGGTGCA 96
 QY 361 GGAGGCTTCCAGAGAGCTGTTGTAATTAGGACCCCAAGCACTGGGAGGGGCTGTGGCT 420
 Db 95 GGAGGCTTCCAGAGAGCTGTTGTAATTAGGACCCCAAGCACTGGGAGGGGCTGTGGCT 36

RESULT 7
 US-09-908-975-16197
 ; Sequence 16197, Application US/09908975
 ; Publication No. US20030165843A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SHOSHAN, Avi
 ; APPLICANT: WASSERMAN, Alon
 ; APPLICANT: MINTZ, Eli
 ; APPLICANT: MINTZ, Liat
 ; APPLICANT: FAIGLER, Simchon
 ; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
 ; FILE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
 ; FILE REFERENCE: 36688-0005
 ; CURRENT APPLICATION NUMBER: US/09/908,975
 ; CURRENT FILING DATE: 2001-07-20
 ; PRIOR APPLICATION NUMBER: US 60/287,724
 ; PRIOR FILING DATE: 2001-05-02
 ; PRIOR APPLICATION NUMBER: US 60/221,607
 ; PRIOR FILING DATE: 2000-07-28
 ; NUMBER OF SEQ ID NOS: 32337
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 16197
 ; LENGTH: 60
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-908-975-16197

Query Match 4.5%; Score 60; DB 10; Length 60;
 Best Local Similarity 100.0%; Pred. No. 1.2e-19;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1169 CTTGGCAAAAGGAATTGTCTAGCCCTACCCCAACCTTCAACTACCAAGATCTGGCCAC 1228
 Db 1 CTTGGCAAAAGGAATTGTCTAGCCCTACCCCAACCTTCAACTACCAAGATCTGGCCAC 60

RESULT 8
 US-10-116-275-299/c
 ; Sequence 299, Application US/10116275
 ; Publication No. US20030211476A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eran Pharmaceutical Technology
 ; APPLICANT: O'Mahony, Daniel J.
 ; APPLICANT: Brayden, David
 ; APPLICANT: Byrne, Daragh
 ; APPLICANT: Landkin, Imelda
 ; APPLICANT: Higgins, Lisa
 ; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
 ; FILE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
 ; FILE REFERENCE: E1067/20087
 ; CURRENT APPLICATION NUMBER: US/10/116,275
 ; CURRENT FILING DATE: 2002-10-04
 ; NUMBER OF SEQ ID NOS: 349
 ; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 299
; LENGTH: 1652
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-116-275-299

Query Match 2.0%; Score 29; DB 16; Length 1652;
Best Local Similarity 100.0%; Pred. No. 0.00059;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1299 ACAGATGTCACAAAAA 1327
Db 1474 ACAGATGTCACAAAAA 1446

RESULT 9

US-10-437-963-80513
; Sequence 80513, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 80513
; LENGTH: 1754
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_80128C.1
US-10-437-963-80513

Query Match 2.0%; Score 27; DB 17; Length 1754;
Best Local Similarity 100.0%; Pred. No. 0.0061;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1301 AGATGTCACAAAAA 1327
Db 1728 AGATGTCACAAAAA 1754

RESULT 10

US-10-087-192-2014/c
; Sequence 2014, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2014
; LENGTH: 248436
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(248436)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-2014

Query Match 2.0%; Score 27; DB 13; Length 248436;
Best Local Similarity 100.0%; Pred. No. 0.0053;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1301 AGATGTCACAAAAA 1327
Db 187916 AGATGTCACAAAAA 187890

RESULT 11

US-09-969-034-4195
; Sequence 4195, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poornima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; TITLE OF INVENTION: Expressed in Cancer Tissue
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4195
; LENGTH: 273
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-034-4195

Query Match 2.0%; Score 26; DB 12; Length 273;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1302 GATGTCACAAAAA 1327
Db 245 GATGTCACAAAAA 270

RESULT 12

US-09-960-352-6255/c
; Sequence 6255, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6255
; LENGTH: 393
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 27-LTB3057-011-Q1-K1-G3
US-09-960-352-6255

```
Query Match          2.0%; Score 26; DB 9; Length 393;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1302 GATGTCACAAAAAAGAAAAA 1327
Db 49 GATGTCACAAAAAAGAAAAA 24

RESULT 13
US-09-960-352-13315/c
; Sequence 13315, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 13315
; LENGTH: 411
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 57-LIB3057-002-Q1-K1-G2
US-09-960-352-13315

Query Match          2.0%; Score 26; DB 9; Length 411;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1302 GATGTCACAAAAAAGAAAAA 1327
Db 48 GATGTCACAAAAAAGAAAAA 23

RESULT 14
US-10-295-027-141
; Sequence 141, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynnne, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1164
; LENGTH: 4702
; TYPE: DNA
; ORGANISM: Homo sapiens

Query Match          2.0%; Score 26; DB 16; Length 4702;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1302 GATGTCACAAAAAAGAAAAA 1327
Db 4660 GATGTCACAAAAAAGAAAAA 4685

RESULT 15
US-10-295-027-1164
; Sequence 1164, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynnne, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1164
; LENGTH: 4702
; TYPE: DNA
; ORGANISM: Homo sapiens
```

US-10-295-027-1164

Query Match 2.0%; Score 26; DB 16; Length 4702;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1302 GATGTCACAAAAA 1327
Db 4660 GATGTCACAAAAA 4685

RESULT 16

US-10-437-963-87284
; Sequence 87284, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 87284
; LENGTH: 6808
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(6808)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_86244C.1
US-10-437-963-87284

Query Match 2.0%; Score 26; DB 17; Length 6808;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1302 GATGTCACAAAAA 1327
Db 6768 GATGTCACAAAAA 6793

RESULT 17

US-10-085-783A-47925
; Sequence 47925, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 47925
; LENGTH: 136
; TYPE: DNA

; ORGANISM: Human
US-10-085-783A-47925

Query Match 1.9%; Score 25; DB 13; Length 136;
Best Local Similarity 100.0%; Pred. No. 0.068;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 104 ATGTCACAAAAA 128

RESULT 18

US-10-242-535A-47925
; Sequence 47925, Application US/10242535A
; Publication No. US2004003663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 47925
; LENGTH: 136
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-47925

Query Match 1.9%; Score 25; DB 16; Length 136;
Best Local Similarity 100.0%; Pred. No. 0.068;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 104 ATGTCACAAAAA 128

RESULT 19

US-10-085-783A-51657
; Sequence 51657, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51657
; LENGTH: 143
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature

```
; LOCATION: (24)..(24)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-51657

Query Match
Best Local Similarity 1.9%; Score 25; DB 13; Length 143;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA... 1327
Db 113 ATGTCACAAAAA... 137

RESULT 20
US-10-242-535A-51657
; Sequence 51657, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 51657
; LENGTH: 143
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (24)..(24)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-51657

Query Match
Best Local Similarity 1.9%; Score 25; DB 16; Length 143;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA... 1327
Db 113 ATGTCACAAAAA... 137

RESULT 21
US-10-085-783A-45661
; Sequence 45661, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 45661

; LENGTH: 156
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-45661

Query Match
Best Local Similarity 1.9%; Score 25; DB 13; Length 156;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA... 1327
Db 125 ATGTCACAAAAA... 149

RESULT 22
US-10-242-535A-45661
; Sequence 45661, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 45661
; LENGTH: 156
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-45661

Query Match
Best Local Similarity 1.9%; Score 25; DB 16; Length 156;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA... 1327
Db 125 ATGTCACAAAAA... 149

RESULT 23
US-09-867-701-9253
; Sequence 9253, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9253
; LENGTH: 184
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-9253

Query Match
Best Local Similarity 1.9%; Score 25; DB 9; Length 184;
Matches 100.0%; Pred. No. 0.068;
Mismatches 0; Indels 0; Gaps 0;
```

```
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1303 ATGTCACAAAAA 1327
Db 141 ATGTCACAAAAA 165

Query Match
US-09-962-436-134/c
; Sequence 134, Application US/09962436
; Patent No. US20020081301A1
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 134
; LENGTH: 191
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-134

Query Match
US-09-954-456-1020/c
; Sequence 1020, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954,456
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1020
```

```
; LENGTH: 191
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1020

Query Match
US-09-880-107-525/c
; Sequence 525, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; PRIOR FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 525
; LENGTH: 191
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA243133
US-09-880-107-525
```

```
Query Match
US-09-867-701-10779
; Sequence 10779, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10779
; LENGTH: 207
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-10779

Query Match
US-09-867-701-10779
; Sequence 10779, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10779
; LENGTH: 207
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-10779
```

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
 Db 164 ATGTCACAAAAA 188

RESULT 28

US-09-880-107-587/c
 ; Sequence 587, Application US/09880107
 ; Patent No. US20020142981A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Horne, Darci T.
 ; APPLICANT: Vockley, Joseph G.
 ; APPLICANT: Scherf, Uwe
 ; APPLICANT: Gene Logic, Inc.
 ; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
 ; FILE REFERENCE: 44921-5028-WO
 ; CURRENT APPLICATION NUMBER: US/09/880,107
 ; CURRENT FILING DATE: 2001-06-14
 ; PRIOR APPLICATION NUMBER: US 60/211,379
 ; PRIOR FILING DATE: 2000-06-14
 ; PRIOR APPLICATION NUMBER: US 60/237,054
 ; PRIOR FILING DATE: 2000-10-02
 ; NUMBER OF SEQ ID NOS: 3950
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 587
 ; LENGTH: 222
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AA258131
 US-09-880-107-587

Query Match 1.9%; Score 25; DB 9; Length 222;
 Best Local Similarity 100.0%; Pred. No. 0.067;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
 Db 39 ATGTCACAAAAA 15

RESULT 29

US-10-424-599-31507/c
 ; Sequence 31507, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J
 ; APPLICANT: Kovalic, David K
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 31507
 ; LENGTH: 256
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_128456C.1
 US-10-424-599-31507

Query Match 1.9%; Score 25; DB 13; Length 256;
 Best Local Similarity 100.0%; Pred. No. 0.067;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
 Db 32 ATGTCACAAAAA 8

RESULT 30

US-10-085-783A-52889
 ; Sequence 52889, Application US/10085783A
 ; Publication No. US20040037841A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ChondroGene Inc.
 ; APPLICANT: Liew, C.C.
 ; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
 ; FILE REFERENCE: 4231/2002
 ; CURRENT APPLICATION NUMBER: US/10/085,783A
 ; CURRENT FILING DATE: 2002-02-28
 ; PRIOR APPLICATION NUMBER: US 60/305,340
 ; PRIOR FILING DATE: 2001-07-13
 ; PRIOR APPLICATION NUMBER: US 60/275,017
 ; PRIOR FILING DATE: 2001-03-12
 ; PRIOR APPLICATION NUMBER: US 60/271,955
 ; PRIOR FILING DATE: 2001-02-28
 ; NUMBER OF SEQ ID NOS: 58994
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 52889
 ; LENGTH: 296
 ; TYPE: DNA
 ; ORGANISM: Human
 ; OTHER INFORMATION: US-10-085-783A-52889

Query Match 1.9%; Score 25; DB 13; Length 296;
 Best Local Similarity 100.0%; Pred. No. 0.067;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
 Db 265 ATGTCACAAAAA 289

RESULT 31

US-10-242-535A-52889
 ; Sequence 52889, Application US/10242535A
 ; Publication No. US20040013663A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ChondroGene Inc.
 ; APPLICANT: Liew, C.C.
 ; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
 ; FILE REFERENCE: 4231/2005
 ; CURRENT APPLICATION NUMBER: US/10/242,535A
 ; CURRENT FILING DATE: 2002-09-12
 ; PRIOR APPLICATION NUMBER: US 10/085,783
 ; PRIOR FILING DATE: 2002-02-28
 ; PRIOR APPLICATION NUMBER: US 60/305,340
 ; PRIOR FILING DATE: 2001-07-13
 ; PRIOR APPLICATION NUMBER: US 60/275,017
 ; PRIOR FILING DATE: 2001-03-12
 ; PRIOR APPLICATION NUMBER: US 60/271,955
 ; PRIOR FILING DATE: 2001-02-28
 ; NUMBER OF SEQ ID NOS: 58994
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 52889
 ; LENGTH: 296
 ; TYPE: DNA
 ; ORGANISM: Human
 ; OTHER INFORMATION: US-10-242-535A-52889

Query Match 1.9%; Score 25; DB 16; Length 296;
 Best Local Similarity 100.0%; Pred. No. 0.067;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
 Db 265 ATGTCACAAAAA 289

RESULT 32

```
US-09-813-358-64/c
; Sequence 64, Application US/09813358
; Patent No. US20020048759A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Pyle, Ruth
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN AND ENDOMETRIAL CANCER
; FILE REFERENCE: 210121.501
; CURRENT APPLICATION NUMBER: US/09/813,358
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 222
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 330
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(330)
; OTHER INFORMATION: n = A,T,C or G
US-09-813-358-64

Query Match      1.9%; Score 25; DB 9; Length 330;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 30 ATGTCACAAAAA 6

RESULT 33
US-09-997-279-64/c
; Sequence 64, Application US/09997279
; Publication No. US20030059781A1
; GENERAL INFORMATION:
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN AND ENDOMETRIAL CANCER
; FILE REFERENCE: 210121.501C1
; CURRENT APPLICATION NUMBER: US/09/997,279
; CURRENT FILING DATE: 2001-11-28
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 330
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(330)
; OTHER INFORMATION: n = A,T,C or G
US-09-997-279-64

Query Match      1.9%; Score 25; DB 10; Length 330;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 30 ATGTCACAAAAA 6

RESULT 34
US-09-922-293-42/c
; Sequence 42, Application US/09922293
; Publication No. US2004012339A1
; GENERAL INFORMATION:
; APPLICANT: Corner, Timothy W.
; APPLICANT: Heck, Gregory R.
```

PRIOR FILING DATE: 1998-05-21
PRIOR APPLICATION NUMBER: US 60/086,184
PRIOR FILING DATE: 1998-05-21
PRIOR APPLICATION NUMBER: US 60/086,183
PRIOR FILING DATE: 1998-05-21
PRIOR APPLICATION NUMBER: US 60/086,188
PRIOR FILING DATE: 1998-05-21
PRIOR APPLICATION NUMBER: US 60/089,524
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: US 60/089,810
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: US 60/089,814
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: US 60/089,808
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: US 60/089,812
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: US 60/089,807
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: US 60/089,806
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: US 60/089,813
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: US 60/089,811
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: US 60/089,793
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: US 60/091,405
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/091,247
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/099,667
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: US 60/099,668
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: US 60/099,670
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: US 60/099,697
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: US 60/100,674
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: US 60/100,673
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: US 60/100,672
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: US 60/100,963
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: US 60/101,131
PRIOR FILING DATE: 1998-09-21
PRIOR APPLICATION NUMBER: US 60/101,132
PRIOR FILING DATE: 1998-09-21
PRIOR APPLICATION NUMBER: US 60/101,130
PRIOR FILING DATE: 1998-09-21
PRIOR APPLICATION NUMBER: US 60/101,508
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: US 60/101,344
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: US 60/101,347
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: US 60/101,343
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: US 60/101,707
PRIOR FILING DATE: 1998-09-25
PRIOR APPLICATION NUMBER: US 60/104,126
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: US 60/104,128
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: US 60/104,127
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: US 60/104,124
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: US 60/109,018
PRIOR FILING DATE: 1998-11-18

PRIOR APPLICATION NUMBER: US 60/108,996
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: US 09/199,129
PRIOR FILING DATE: 1998-11-24
PRIOR APPLICATION NUMBER: US 09/210,297
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: US 60/111,981
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: US 60/113,224
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: US 09/229,413
PRIOR FILING DATE: 1999-01-12
NUMBER OF SEQ ID NOS: 3853
SEQ ID NO 42
LENGTH: 330
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-922-293-42

Query Match 1.9%; Score 25; DB 12; Length 330;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCAAAAAAAAAAAAAAAAAAAAA 1327
|||||
Db 101 ATGTCAAAAAAAAAAAAAAAAAAAAA 77

RESULT 35

US-09-960-352-1456/c
Sequence 1456, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Mengbing
APPLICANT: Wyatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 1456
LENGTH: 354
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 07-LIB3057-005-Q1-K1-B3
US-09-960-352-1456

Query Match 1.9%; Score 25; DB 9; Length 354;
Best Local Similarity 100.0%; Pred. No. 0.067;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1302 GATGTCAAAAAAAAAAAAAAAAAAAA 1326
|||||
Db 39 GATGTCAAAAAAAAAAAAAAAAAAAA 15

RESULT 36

US-10-424-599-52765
Sequence 52765, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: Ia Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28


```
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 52765
; LENGTH: 358
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_1865C.1
US-10-424-599-52765

Query Match
Best Local Similarity 100.0%; Score 25; DB 13; Length 358;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 22 ATGTCACAAAAA 46

RESULT 37
US-10-437-963-88420
; Sequence 88420, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 88420
; LENGTH: 382
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(382)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_87273C.1
US-10-437-963-88420

Query Match
Best Local Similarity 100.0%; Score 25; DB 17; Length 382;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 355 ATGTCACAAAAA 379

RESULT 38
US-10-175-184A-31
; Sequence 31, Application US/10175184A
; Publication No. US20040038292A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc.
; APPLICANT: Burslem, Martyn Frank
; APPLICANT: Johnson, Claire Michelle
; APPLICANT: Cooper, Lisa
; APPLICANT: Martin, Paul
; TITLE OF INVENTION: WOUND HEALING BIOMARKERS
; FILE REFERENCE: PC22024GLK
; CURRENT APPLICATION NUMBER: US/10/175,184A
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/305,346
```

```
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: GB 0114869.1
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31
; LENGTH: 390
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-175-184A-31

Query Match
Best Local Similarity 100.0%; Score 25; DB 13; Length 390;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 356 ATGTCACAAAAA 380

RESULT 39
US-09-918-995-7699
; Sequence 7699, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7699
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-7699

Query Match
Best Local Similarity 100.0%; Score 25; DB 10; Length 430;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 405 ATGTCACAAAAA 429

RESULT 40
US-09-960-352-8587/c
; Sequence 8587, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Rao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 8587
; LENGTH: 464
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 37-LIB3057-005-Q1-K1-B2
US-09-960-352-8587

Query Match
Best Local Similarity 100.0%; Score 25; DB 9; Length 464;
```

```
Best Local Similarity 100.0%; Pred. No. 0.066;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1302 GATGTCACAAAAA 1326
Db 42 GATGTCACAAAAA 18

RESULT 41
US-10-198-846-1221/c
; Sequence 1221, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1221
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 5, 7, 14, 116, 200, 230, 243, 245, 252, 294, 312, 340, 349,
; LOCATION: 355, 446, 448, 452, 462
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-1221

Query Match 1.9%; Score 25; DB 15; Length 467;
Best Local Similarity 100.0%; Pred. No. 0.066;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 99 ATGTCACAAAAA 75

RESULT 42
US-10-102-524-36/c
; Sequence 36, Application US/10102524
; Publication No. US2003010943A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Gaiger, Alexander
; APPLICANT: Gordon, Brian
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF KIDNEY CANCER
; FILE REFERENCE: 210121.572
; CURRENT APPLICATION NUMBER: US/10102,524
; CURRENT FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 1863
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 64, 100, 111, 120, 148, 163, 298, 327, 338, 430, 455, 462,
; LOCATION: 471
; OTHER INFORMATION: n = A,T,C or G
```

```
US-10-102-524-36

Query Match 1.9%; Score 25; DB 15; Length 483;
Best Local Similarity 100.0%; Pred. No. 0.066;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 39 ATGTCACAAAAA 15

RESULT 43
US-10-437-963-85418
; Sequence 85418, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 85418
; LENGTH: 483
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(483)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: PAT_MRT4530_84559C.1
US-10-437-963-85418

Query Match 1.9%; Score 25; DB 17; Length 483;
Best Local Similarity 100.0%; Pred. No. 0.066;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 421 ATGTCACAAAAA 445

RESULT 44
US-10-027-632-95705
; Sequence 95705, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
```

```
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 95705
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-95705

Query Match
Best Local Similarity 1.9%; Score 25; DB 13; Length 486;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 143 ATGTCACAAAAA 167

RESULT 45
US-10-027-632-305902
; Sequence 305902, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 305902
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-305902

Query Match
Best Local Similarity 1.9%; Score 25; DB 13; Length 486;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 143 ATGTCACAAAAA 167

RESULT 46
US-10-027-632-95705
; Sequence 95705, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
```

```
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 95705
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-95705
```

```
Query Match
Best Local Similarity 1.9%; Score 25; DB 16; Length 486;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 143 ATGTCACAAAAA 167
```

```
RESULT 47
US-10-027-632-305902
; Sequence 305902, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 305902
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-305902
```

```
Query Match
Best Local Similarity 1.9%; Score 25; DB 16; Length 486;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
Db 143 ATGTCACAAAAA 167
```

Fri Sep 17 10:18:25 2004

RESULT 48
US-10-437-963-19508/c
; Sequence 19508, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 19508
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_24962C.1
US-10-437-963-19508

Query Match 1.9%; Score 25; DB 17; Length 492;
Best Local Similarity 100.0%; Pred. No. 0.066;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCACAAAAA 1327
Db 35 ATGTCACAAAAA 11

RESULT 49
US-10-029-386-1201/c
; Sequence 1201, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEWICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 1201
; LENGTH: 503
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR22 30.0
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
; OTHER INFORMATION: NT HIT: M23442.1, EVALUUE 3.10e-02
US-10-029-386-1201

Query Match 1.9%; Score 25; DB 15; Length 503;
Best Local Similarity 100.0%; Pred. No. 0.066;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCACAAAAA 1327
Db 230 ATGTCACAAAAA 206

RESULT 50
US-09-925-301-787
; Sequence 787, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 787
; LENGTH: 527
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (492)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-787

Query Match 1.9%; Score 25; DB 9; Length 527;
Best Local Similarity 100.0%; Pred. No. 0.066;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1303 ATGTCACAAAAA 1327
Db 501 ATGTCACAAAAA 525

Search completed: September 17, 2004, 06:45:02
Job time : 3730 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 17, 2004, 02:40:43 ; Search time 123 Seconds
(without alignments)
5987.153 Million cell updates/sec

Title: US-09-927-091-3_COPY_2500_3826

Perfect score: 1327

Sequence: 1 tccacagtgtgcacagtag.....aaaaaaaaaaaaaaaa 1327

Scoring table: OLIGO NUC

Gapop_60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database : Issued Patents NA:*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	1.9	469	US-08-468-347-23	Sequence 23, Appl
2	25	1.9	469	US-08-226-264-25	Sequence 25, Appl
3	25	1.9	469	US-08-467-389-23	Sequence 23, Appl
4	25	1.9	469	US-08-779-379-23	Sequence 23, Appl
5	25	1.9	469	US-08-469-219-23	Sequence 23, Appl
6	25	1.9	469	US-09-228-152-23	Sequence 23, Appl
7	25	1.9	729	US-08-447-010-1	Sequence 1, Appl
8	25	1.9	868	US-08-889-502-20	Sequence 20, Appl
C 9	25	1.9	98844	US-09-791-211-10	Sequence 10, Appl
C 10	24	1.8	40	US-09-306-290-26	Sequence 26, Appl
C 11	24	1.8	555	US-07-449-285A-15	Sequence 15, Appl
C 12	24	1.8	1196	US-07-959-509-4	Sequence 4, Appl
13	24	1.8	1332	US-09-333-423-1	Sequence 1, Appl
14	24	1.8	1392	US-08-957-351-1	Sequence 1, Appl
15	24	1.8	1478	US-09-216-393B-7	Sequence 7, Appl
16	24	1.8	1496	US-09-712-529-1	Sequence 1, Appl
17	24	1.8	1628	US-08-883-515-3	Sequence 3, Appl
18	24	1.8	1705	US-09-205-258-216	Sequence 216, Appl
19	24	1.8	2017	US-09-436-983-1	Sequence 1, Appl
20	24	1.8	2193	US-09-427-261-2	Sequence 2, Appl
21	24	1.8	2193	US-09-427-261-3	Sequence 2, Appl
22	24	1.8	3001	US-09-539-333D-222	Sequence 222, Appl
23	24	1.8	3848	US-09-112-096-28	Sequence 28, Appl
24	24	1.8	3952	US-08-381-691-16	Sequence 16, Appl
25	24	1.8	5668	US-09-112-096-14	Sequence 14, Appl
26	24	1.8	5668	US-09-636-215-777	Sequence 777, Appl
27	24	1.8	5668	US-09-685-166A-777	Sequence 777, Appl

C 28	24	1.8	21234	4	US-09-810-671-3	Sequence 3, Appl
C 29	24	1.8	21234	4	US-10-109-854-3	Sequence 3, Appl
C 30	24	1.8	32042	4	US-09-245-281-44	Sequence 44, Appl
C 31	24	1.8	32042	4	US-09-340-620A-63	Sequence 63, Appl
C 32	24	1.8	50000	4	US-09-146-053-3	Sequence 3, Appl
C 33	24	1.8	786431	4	US-09-751-389-3	Sequence 3, Appl
C 34	23	1.7	40	3	US-09-306-290-15	Sequence 15, Appl
35	23	1.7	67	4	US-09-621-976-10956	Sequence 10956, A
36	23	1.7	97	4	US-09-621-976-9086	Sequence 9086, Ap
37	23	1.7	307	4	US-09-091-725-28	Sequence 28, Appl
38	23	1.7	350	4	US-09-621-976-15342	Sequence 15342, A
39	23	1.7	433	1	US-07-987-272A-13	Sequence 13, Appl
40	23	1.7	567	3	US-08-561-168-6	Sequence 6, Appl
41	23	1.7	627	3	US-09-385-982-4	Sequence 4, Appl
42	23	1.7	989	2	US-08-874-460-1	Sequence 1, Appl
43	23	1.7	989	4	US-09-272-162-1	Sequence 1, Appl
44	23	1.7	1005	4	US-09-396-149-17	Sequence 5, Appl
45	23	1.7	1087	4	US-09-647-224A-5	Sequence 17, Appl
46	23	1.7	1108	1	US-08-036-555B-135	Sequence 135, App
47	23	1.7	1108	1	US-08-469-569-135	Sequence 135, App
48	23	1.7	1108	1	US-08-249-322A-135	Sequence 135, App
49	23	1.7	1108	1	US-08-469-526A-135	Sequence 135, App
50	23	1.7	1108	2	US-08-734-591A-135	Sequence 135, App
51	23	1.7	1108	2	US-08-469-660-135	Sequence 135, App
52	23	1.7	1108	3	US-08-341-018-5	Sequence 5, Appl
53	23	1.7	1108	3	US-08-470-335-135	Sequence 135, App
54	23	1.7	1108	3	US-08-735-021-135	Sequence 135, App
55	23	1.7	1108	3	US-08-734-664A-135	Sequence 135, App
56	23	1.7	1108	3	US-08-470-339-135	Sequence 135, App
57	23	1.7	1108	4	US-08-467-602-135	Sequence 135, App
58	23	1.7	1108	5	PCT-US94-05083C-131	Sequence 131, App
59	23	1.7	1108	5	PCT-US95-06846A-135	Sequence 135, App
60	23	1.7	1193	1	US-08-036-555B-134	Sequence 134, App
61	23	1.7	1193	1	US-08-469-569-134	Sequence 134, App
62	23	1.7	1193	1	US-08-249-322A-134	Sequence 134, App
63	23	1.7	1193	1	US-08-469-526A-134	Sequence 134, App
64	23	1.7	1193	2	US-08-734-591A-134	Sequence 134, App
65	23	1.7	1193	2	US-08-469-660-134	Sequence 134, App
66	23	1.7	1193	3	US-08-341-018-3	Sequence 3, Appl
67	23	1.7	1193	3	US-08-470-335-134	Sequence 134, App
68	23	1.7	1193	3	US-08-735-021-134	Sequence 134, App
69	23	1.7	1193	3	US-08-734-664A-134	Sequence 134, App
70	23	1.7	1193	3	US-08-470-339-134	Sequence 134, App
71	23	1.7	1193	4	US-08-467-602-134	Sequence 134, App
72	23	1.7	1193	5	PCT-US94-05083C-130	Sequence 130, App
73	23	1.7	1193	5	PCT-US95-06846A-134	Sequence 134, App
74	23	1.7	1209	4	US-09-614-912-61	Sequence 61, Appl
75	23	1.7	1291	4	US-08-555-755C-5	Sequence 5, Appl
76	23	1.7	1364	2	US-08-815-718-1	Sequence 1, Appl
77	23	1.7	1509	4	US-09-149-476-173	Sequence 173, App
78	23	1.7	1559	4	US-09-482-273-56	Sequence 56, Appl
79	23	1.7	1575	4	US-09-833-381-1159	Sequence 1159, Ap
C 80	23	1.7	1703	3	US-09-135-021-77	Sequence 77, Appl
C 81	23	1.7	1703	3	US-09-135-020-3	Sequence 3, Appl
C 82	23	1.7	1703	3	US-09-135-010A-3	Sequence 3, Appl
C 83	23	1.7	1703	4	US-09-444-871-3	Sequence 3, Appl
C 84	23	1.7	1703	4	US-09-597-735-3	Sequence 3, Appl
C 85	23	1.7	1703	4	US-09-444-285-3	Sequence 3, Appl
C 86	23	1.7	1703	4	US-09-597-733-3	Sequence 3, Appl
C 87	23	1.7	1703	4	US-09-597-731-3	Sequence 3, Appl
C 88	23	1.7	1741	1	US-08-565-655-5	Sequence 5, Appl
89	23	1.7	1769	4	US-09-333-583A-5	Sequence 5, Appl
90	23	1.7	1769	4	US-09-503-250-1	Sequence 1, Appl
91	23	1.7	1769	4	US-09-016-434-1072	Sequence 1072, Ap
92	23	1.7	1769	4	US-09-023-685-888	Sequence 888, App
C 93	23	1.7	1821	3	US-09-056-105-23	Sequence 23, Appl
C 94	23	1.7	1896	1	US-08-253-503-1	Sequence 1, Appl
C 95	23	1.7	1896	1	US-08-796-883-1	Sequence 1, Appl
C 96	23	1.7	1896	2	US-08-611-273B-1	Sequence 1, Appl
C 97	23	1.7	1896	2	US-08-531-864-1	Sequence 1, Appl
C 98	23	1.7	1896	2	US-08-373-636C-1	Sequence 1, Appl
C 99	23	1.7	1896	3	US-08-603-506A-1	Sequence 1, Appl
C 100	23	1.7	1896	3	US-09-266-294-1	Sequence 1, Appl

c 101	23	1.7	1896	3	US-09-179-281-1	Sequence 1, Appli
c 102	23	1.7	1896	3	US-09-056-105-25	Sequence 25, Appli
c 103	23	1.7	1969	2	US-08-541-033A-7	Sequence 7, Appli
c 104	23	1.7	1969	2	US-08-828-451-7	Sequence 7, Appli
c 105	23	1.7	2026	1	US-08-503-133A-1	Sequence 1, Appli
c 106	23	1.7	2026	2	US-08-576-775A-1	Sequence 1, Appli
c 107	23	1.7	2026	2	US-08-972-499-1	Sequence 1, Appli
c 108	23	1.7	2026	3	US-08-899-545-1	Sequence 1, Appli
c 109	23	1.7	2053	4	US-09-833-381-1403	Sequence 1403, Ap
c 110	23	1.7	2096	2	US-08-541-033A-19	Sequence 19, Appli
c 111	23	1.7	2096	2	US-08-828-451-19	Sequence 19, Appli
c 112	23	1.7	2099	2	US-08-541-033A-3	Sequence 3, Appli
c 113	23	1.7	2099	2	US-08-828-451-3	Sequence 3, Appli
c 114	23	1.7	2103	2	US-08-897-340-2	Sequence 2, Appli
c 115	23	1.7	2103	3	US-09-252-329-2	Sequence 2, Appli
c 116	23	1.7	2119	2	US-08-381-691-17	Sequence 17, Appli
c 117	23	1.7	2137	2	US-08-541-033A-18	Sequence 18, Appli
c 118	23	1.7	2137	2	US-08-828-451-18	Sequence 18, Appli
c 119	23	1.7	2140	2	US-08-541-033A-1	Sequence 1, Appli
c 120	23	1.7	2140	2	US-08-828-451-1	Sequence 1, Appli
c 121	23	1.7	2198	2	US-08-755-728-2	Sequence 2, Appli
c 122	23	1.7	2198	2	US-08-974-655-2	Sequence 2, Appli
c 123	23	1.7	2198	3	US-09-283-011-2	Sequence 2, Appli
c 124	23	1.7	2209	4	US-09-807-258-25	Sequence 25, Appli
c 125	23	1.7	2335	3	US-09-387-574-9	Sequence 9, Appli
c 126	23	1.7	2335	4	US-09-668-096-9	Sequence 9, Appli
c 127	23	1.7	2403	4	US-09-614-912-99	Sequence 99, Appli
c 128	23	1.7	2420	5	PCT-US93-00227-1	Sequence 1, Appli
c 129	23	1.7	2759	4	US-09-144-367-1	Sequence 1, Appli
c 130	23	1.7	2929	4	US-09-705-299-10	Sequence 10, Appli
c 131	23	1.7	3238	4	US-08-123-934A-5	Sequence 5, Appli
c 132	23	1.7	3238	4	US-09-874-628-5	Sequence 5, Appli
c 133	23	1.7	3238	5	PCT-US94-10080-5	Sequence 5, Appli
c 134	23	1.7	3268	4	US-09-688-188B-1	Sequence 1, Appli
c 135	23	1.7	3268	4	US-09-291-417D-1	Sequence 1, Appli
c 136	23	1.7	3300	4	US-09-336-643A-82	Sequence 82, Appli
c 137	23	1.7	3377	4	US-09-819-989-1	Sequence 1, Appli
c 138	23	1.7	3377	4	US-10-273-992-1	Sequence 1, Appli
c 139	23	1.7	3441	4	US-09-866-028-6	Sequence 6, Appli
c 140	23	1.7	3946	3	US-09-083-351-1	Sequence 1, Appli
c 141	23	1.7	3946	3	US-09-083-352-1	Sequence 1, Appli
c 142	23	1.7	4113	4	US-09-785-381-2	Sequence 2, Appli
c 143	23	1.7	4843	3	US-08-986-485-1	Sequence 1, Appli
c 144	23	1.7	4874	4	US-09-187-330-2	Sequence 2, Appli
c 145	23	1.7	6182	4	US-10-204-708-88	Sequence 88, Appli
c 146	23	1.7	6405	4	US-09-281-481A-18	Sequence 18, Appli
c 147	23	1.7	6644	4	US-08-873-435B-5	Sequence 5, Appli
c 148	23	1.7	8201	1	US-08-253-155A-9	Sequence 9, Appli
c 149	23	1.7	9278	1	US-08-243-542-9	Sequence 9, Appli
c 150	23	1.7	9378	1	US-08-477-407-9	Sequence 9, Appli

ALIGNMENTS

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,347
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/225,442
FILING DATE: 08-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0317/43020-A-JPW/EAB
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-977-9550
TELEFAX: 212-664-0525
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-468-347-23

Query Match          1.9%; Score 25; DB 1; Length 469;
Best Local Similarity 100.0%; Pred.No. 0.053;
Matches      25; Conservative    0; Mismatches   0; Indels

Qy      1303 ATGTCAAAAAAAAAAAAAAAAAAAAAA 1327
        |||
Db      445 ATGTCAAAAAAAAAAAAAAAAAAAAAA 469

RESULT 2
US-08-226-264-25
Sequence 25, Application US/08226264
Patent No. 5801017
GENERAL INFORMATION:
APPLICANT: Werber, Moshe M.
APPLICANT: Zeelon, Elisha P.
APPLICANT: Levanon, Avigdor
APPLICANT: Guy, Rachel
APPLICANT: Goldlust, Arie
APPLICANT: Rigbi, Meir
APPLICANT: Panet, Amos
APPLICANT: Fischer, Meir
TITLE OF INVENTION: PRODUCTION OF RECOMBINANT FACTOR XA
OF INVENTION: INHIBITORS OF LEECH HIRUDO MEDICINALIS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/226,264
FILING DATE: 08-APR-94
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
```

REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40017-A/JPW/GJG
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
TELEX:
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-226-264-25

Query Match 1.9%; Score 25; DB 1; Length 469;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
DB 445 ATGTCACAAAAA 469

RESULT 3

US-08-467-389-23
Sequence 23, Application US/08467389
Patent No. 5824641
GENERAL INFORMATION:
APPLICANT: Zeelon, Elisha P.
APPLICANT: Werber, Moshe M.
APPLICANT: Levanon, Avigdor
TITLE OF INVENTION: NOVEL POLYPEPTIDE HAVING FACTOR Xa
TITLE OF INVENTION: INHIBITORY ACTIVITY
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
ZIP: 10112

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,389
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/225,442
FILING DATE: 08-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0317/43020-A/JPW/EAB
TELEPHONE: 212-977-9550
TELEFAX: 212-664-0525

INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-467-389-23

Query Match 1.9%; Score 25; DB 1; Length 469;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
DB 445 ATGTCACAAAAA 469

RESULT 4

US-08-779-379-23
Sequence 23, Application US/08779379
Patent No. 5858970
GENERAL INFORMATION:
APPLICANT: Zeelon, Elisha P.
APPLICANT: Werber, Moshe M.
APPLICANT: Levanon, Avigdor
TITLE OF INVENTION: NOVEL POLYPEPTIDE HAVING FACTOR Xa
TITLE OF INVENTION: INHIBITORY ACTIVITY
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
ZIP: 10112

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/779,379
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/225,442
FILING DATE: 08-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0317/43020-A/JPW/EAB
TELEPHONE: 212-977-9550
TELEFAX: 212-664-0525

INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-779-379-23

Query Match 1.9%; Score 25; DB 2; Length 469;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
DB 445 ATGTCACAAAAA 469

RESULT 5

US-08-469-219-23
Sequence 23, Application US/08469219
Patent No. 5863534
GENERAL INFORMATION:

APPLICANT: Zeelon, Elisha P.
APPLICANT: Werber, Moshe M.
APPLICANT: Levanon, Avigdor
TITLE OF INVENTION: NOVEL POLYPEPTIDE HAVING FACTOR Xa
TITLE OF INVENTION: INHIBITORY ACTIVITY
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,219
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/225,442
FILING DATE: 08-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0317/43020-A/JPM/EAB
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-977-9550
TELEFAX: 212-664-0525
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-469-219-23

Query Match 1.9%; Score 25; DB 2; Length 469;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCAAAAAAAAAAAAAAAAAAAAA 1327
Db 445 ATGTCAAAAAAAAAAAAAAAAAAAAA 469

RESULT 6
US-09-228-152-23
Sequence 23, Application US/09228152
Patent No. 6211341
GENERAL INFORMATION:
APPLICANT: Zeelon, Elisha P.
APPLICANT: Werber, Moshe M.
APPLICANT: Levanon, Avigdor
TITLE OF INVENTION: NOVEL POLYPEPTIDE HAVING FACTOR Xa INHIBITORY ACTIVITY
FILE REFERENCE: 43020aya
CURRENT APPLICATION NUMBER: US/09/228,152
CURRENT FILING DATE: 1999-01-11
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 23
LENGTH: 469
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
OTHER INFORMATION: of clone pSP65-XaI-4.

US-09-228-152-23
Query Match 1.9%; Score 25; DB 3; Length 469;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCAAAAAAAAAAAAAAAAAAAAA 1327
Db 445 ATGTCAAAAAAAAAAAAAAAAAAAAA 469

RESULT 7
US-08-447-010-1
Sequence 1, Application US/08447010
Patent No. 5770718
GENERAL INFORMATION:
APPLICANT: MOFFATT, BARBARA
TITLE OF INVENTION: GENE FOR APRT FROM PLANT TISSUE
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 330 University Avenue, Suite 701
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/447,010
FILING DATE: 22-MAY-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/230,695
FILING DATE: 21-APR-1994
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/888,132
FILING DATE: 26-MAY-1992
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, MICHAEL I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1811-183 MIS:vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
TELEX: 065-24567 SIMBAS
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 729 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(18..569)
US-08-447-010-1

Query Match 1.9%; Score 25; DB 1; Length 729;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCAAAAAAAAAAAAAAAAAAAAA 1327
Db 705 ATGTCAAAAAAAAAAAAAAAAAAAAA 729

RESULT 8

US-08-889-502-20
; Sequence 20, Application US/08889502
; Patent No. 6066726
; GENERAL INFORMATION:
; APPLICANT: Parb, David H
; APPLICANT: Russek, Shelley J
; TITLE OF INVENTION: GENE THERAPY VECTOR WITH TISSUE
; TITLE OF INVENTION: SPECIFICITY
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kevin M. Farrell
; STREET: P.O. Box 999
; CITY: York Harbor
; STATE: ME
; COUNTRY: USA
; ZIP: 03911
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/889,502
; FILING DATE: 08-JUL-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Farrell, Kevin M
; REGISTRATION NUMBER: 35,505
; REFERENCE/DOCKET NUMBER: 0146-2008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (207) 363-0558
; TELEFAX: (207) 363-0528
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 868 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-889-502-20

Query Match 1.9%; Score 25; DB 3; Length 868;
Best Local Similarity 100.0%; Pred. No. 0.053;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
|||||
Db 440 ATGTCACAAAAA 464

RESULT 9
US-09-791-211-10/c
; Sequence 10, Application US/09791211
; Patent No. 6448080
; GENERAL INFORMATION:
; APPLICANT: Donna T. Ward
; APPLICANT: Andrew T. Watt
; TITLE OF INVENTION: ANTISENSE MODULATION OF WRN EXPRESSION
; FILE REFERENCE: RTS-0205
; CURRENT APPLICATION NUMBER: US/09/791,211
; CURRENT FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 10
; LENGTH: 98844
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 24962
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 64383
; OTHER INFORMATION: unknown

; NAME/KEY: unsure
; LOCATION: 65468
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 65469
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 65470
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 65471
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 87130
; OTHER INFORMATION: unknown
; NAME/KEY: unsure
; LOCATION: 89049
; OTHER INFORMATION: unknown
; OTHER INFORMATION:
US-09-791-211-10

Query Match 1.9%; Score 25; DB 4; Length 98844;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1303 ATGTCACAAAAA 1327
|||||
Db 95450 ATGTCACAAAAA 95426

RESULT 10
US-09-306-290-26/c
; Sequence 26, Application US/09306290
; Patent No. 6221635
; GENERAL INFORMATION:
; APPLICANT: Rovera, Giovanni
; APPLICANT: Mukhopadhyay, Sunil
; TITLE OF INVENTION: METHODS FOR SOLID-PHASE AMPLIFICATION OF DNA TEMPLATE
; TITLE OF INVENTION: (SPADT) USING MULTIPLE ARRAYS
; FILE REFERENCE: 09924-10
; CURRENT APPLICATION NUMBER: US/09/306,290
; CURRENT FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 26
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer p41
; OTHER INFORMATION: FH440
US-09-306-290-26

Query Match 1.8%; Score 24; DB 3; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1304 TGTCAAAAAA 1327
|||||
Db 24 TGTCAAAAAA 1

RESULT 11
US-09-449-285A-15/c
; Sequence 15, Application US/09449285A
; Patent No. 6313280
; GENERAL INFORMATION:
; APPLICANT: Vlaams Interuniversitair Instituut voor Biotechnologie
; TITLE OF INVENTION: SMAD-INTERACTING POLYPEPTIDES AND THEIR USE
; FILE REFERENCE: 2676-4232US
; CURRENT APPLICATION NUMBER: US/09/449,285A
; CURRENT FILING DATE: 1999-11-24
; PRIOR APPLICATION NUMBER: PCT/EP98/03193

;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 97201645.5
;; PRIOR FILING DATE: 1997-06-02
;; NUMBER OF SEQ ID NOS: 27
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO 15
;; LENGTH: 555
;; TYPE: DNA
;; ORGANISM: Mus musculus
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (382)..(555)
;; OTHER INFORMATION: n can be any nucleotide
US-09-449-285A-15

Query Match 1.8%; Score 24; DB 4; Length 555;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1304 TGTCAAAAAAAAAAAAAAAAAAAAA 1327
Db 24 TGTCAAAAAAAAAAAAAAAAAAAAA 1

RESULT 12

US-07-959-509-4
;; Sequence 4, Application US/07959509
;; Patent No. 6001560
;; GENERAL INFORMATION:
;; APPLICANT: Lonial, Herinder
;; APPLICANT: Narula, Satwant
;; APPLICANT: Zavodny, Paul
;; TITLE OF INVENTION: Human Gamma Interferon Antagonist/Agonist Screen
;; NUMBER OF SEQUENCES: 4
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Schering-Plough Corporation
;; STREET: One Giralda Farms
;; CITY: Madison
;; STATE: New Jersey
;; COUNTRY: USA
;; ZIP: 07940

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 6.0.5
SOFTWARE: Microsoft Word 4.00B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/959,509
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/616,621
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Dulak, No. 6001560man C.
REGISTRATION NUMBER: 31,608
REFERENCE/DOCKET NUMBER: JBO166

TELECOMMUNICATION INFORMATION:
TELEPHONE: 201 822 7375
TELEFAX: 201 822 7039
TELEX: 219165

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:

LENGTH: 1196 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

US-07-959-509-4

Query Match 1.8%; Score 24; DB 3; Length 1196;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1304 TGTCAAAAAAAAAAAAAAAAAAAAA 1327
Db 327 TGTCAAAAAAAAAAAAAAAAAAAAA 350

RESULT 13

US-09-333-423-1
;; Sequence 1, Application US/09333423
;; Patent No. 6265636
;; GENERAL INFORMATION:
;; APPLICANT: Randall, Douglas
;; APPLICANT: Thelen, Jay
;; APPLICANT: Miernyk, Jan
;; APPLICANT: Muszynski, Michael
;; APPLICANT: Sewalt, Vincent
;; TITLE OF INVENTION: Pyruvate Dehydrogenase Kinase
;; TITLE OF INVENTION: Polynucleotides, Polypeptides and Uses Thereof
;; FILE REFERENCE: 0818
;; CURRENT APPLICATION NUMBER: US/09/333,423
;; CURRENT FILING DATE: 1999-06-15
;; EARLIER APPLICATION NUMBER: 60/089,998
;; EARLIER FILING DATE: 1998-06-19
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 1
;; LENGTH: 1332
;; TYPE: DNA
;; ORGANISM: Zea mays
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (55)...(1095)
US-09-333-423-1

Query Match 1.8%; Score 24; DB 3; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1304 TGTCAAAAAAAAAAAAAAAAAAAAA 1327
Db 1279 TGTCAAAAAAAAAAAAAAAAAAAAA 1302

RESULT 14

US-08-957-351-1
;; Sequence 1, Application US/08957351
;; Patent No. 6306586
;; GENERAL INFORMATION:
;; APPLICANT: Semina, Elena
;; APPLICANT: Murray, Jeffrey C.
;; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE
;; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF CATARACTS
;; NUMBER OF SEQUENCES: 33
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
;; STREET: One Post Office Square
;; CITY: Boston
;; STATE: MA
;; COUNTRY: USA
;; ZIP: 02109-2170

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/957,351
FILING DATE: 24-OCT-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Arnold, Beth E.
REGISTRATION NUMBER: 35,430
REFERENCE/DOCKET NUMBER: UIA-024.01
TELECOMMUNICATION INFORMATION:

```
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1392 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-957-351-1

Query Match      1.8%; Score 24; DB 4; Length 1392;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1304 TGTCAAAAAAAAAAAAAAAAAAAAAA 1327
Db 1369 TGTCAAAAAAAAAAAAAAAAAAAAAA 1392

RESULT 15
US-09-216-393B-7
; Sequence 7, Application US/09216393B
; Patent No. 6514694
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393B
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/994,825
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1478
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(1161)
; OTHER INFORMATION:
US-09-216-393B-7

Query Match      1.8%; Score 24; DB 4; Length 1478;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1304 TGTCAAAAAAAAAAAAAAAAAAAAAA 1327
Db 1427 TGTCAAAAAAAAAAAAAAAAAAAAAA 1450

RESULT 16
US-09-712-529-1
; Sequence 1, Application US/09712529
; Patent No. 6485938
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Bishop, Paul D.
; APPLICANT: Whitmore, Theodore E.
; APPLICANT: Thompson, Penny P.
; TITLE OF INVENTION: Human Zven Proteins
; FILE REFERENCE: 99-81
; CURRENT APPLICATION NUMBER: US/09/712,529
; CURRENT FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1496
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: (66)...(389)
US-09-712-529-1

Query Match      1.8%; Score 24; DB 4; Length 1496;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1304 TGTCAAAAAAAAAAAAAAAAAAAAAA 1327
Db 1459 TGTCAAAAAAAAAAAAAAAAAAAAAA 1482

RESULT 17
US-08-883-515-3
; Sequence 3, Application US/08883515
; Patent No. 5981836
; GENERAL INFORMATION:
; APPLICANT: Osteryoung, Katherine W
; TITLE OF INVENTION: PLANT CHLOROPLAST DIVISION GENES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/883,515
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27,386
; REFERENCE/DOCKET NUMBER: 920905.90016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1628 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..1316
; US-08-883-515-3

Query Match      1.8%; Score 24; DB 2; Length 1628;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1304 TGTCAAAAAAAAAAAAAAAAAAAAAA 1327
Db 1604 TGTCAAAAAAAAAAAAAAAAAAAAAA 1627

RESULT 18
US-09-205-258-216
; Sequence 216, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: PZ007P1
```

```

; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877

; Query Match 1.8%; Score 24; DB 4; Length 1705;
; Best Local Similarity 100.0%; Pred.No. 0.15;
; Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1304 TGTCACAAAAA 1327
Db 1647 TGTCACAAAAA 1670

RESULT 19
US-09-436-983-1
; Sequence 1, Application US/09436983
; Patent No. 6294343
; GENERAL INFORMATION:
; APPLICANT: Mack, David
; APPLICANT: Gish, Kurt C
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING COLORECTAL CANCER,
; TITLE OF INVENTION: COMPOSITIONS, AND METHODS OF SCREENING FOR COLORECTAL
; TITLE OF INVENTION: CANCER MODULATORS
; FILE REFERENCE: A-68431/RMS/DAV
; CURRENT APPLICATION NUMBER: US/09/436,983
; CURRENT FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2017
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-436-983-1
; Query Match 1.8%; Score 24; DB 3; Length 2017;
; Best Local Similarity 100.0%; Pred.No. 0.15;
; Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1304 TGTCACAAAAA 1327
Db 1981 TGTCACAAAAA 2004

RESULT 20
US-09-427-261-2
; Sequence 2, Application US/09427261A
; Patent No. 6368811
; GENERAL INFORMATION:
; APPLICANT: Grootjans, Jan
; APPLICANT: Zimmerman, Pascale
; APPLICANT: David, Guido
```

```
/ TITLE OF INVENTION: Syndecan Interacting Proteins and the Use Thereof
/ FILE REFERENCE: 2676-4206US
/ CURRENT APPLICATION NUMBER: US/09/427,261A
/ CURRENT FILING DATE: 1999-10-25
/ NUMBER OF SEQ ID NOS: 3
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 2193
/ TYPE: DNA
/ ORGANISM: Unknown Organism
/ FEATURE:
/ OTHER INFORMATION: Description of Unknown Organism:Syndecan DNA
US-09-427-261-2

Query Match
Best Local Similarity 100.0%; Score 24; DB 4; Length 2193;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1304 TGTCAAAAAAAAAAAAAAAAAAAAAA 1327
Db 2159 TGTCAAAAAAAAAAAAAAAAAAAAAA 2182

RESULT 21
US-09-427-261-3
/ Sequence 3, Application US/09427261A
/ Patent No. 6368811
/ GENERAL INFORMATION:
/ APPLICANT: Grootjans, Jan
/ APPLICANT: Zimmerman, Pascale
/ APPLICANT: David, Guido
/ TITLE OF INVENTION: Syndecan Interacting Proteins and the Use Thereof
/ FILE REFERENCE: 2676-4206US
/ CURRENT APPLICATION NUMBER: US/09/427,261A
/ CURRENT FILING DATE: 1999-10-25
/ NUMBER OF SEQ ID NOS: 3
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 3
/ LENGTH: 2193
/ TYPE: DNA
/ ORGANISM: Unknown Organism
/ FEATURE:
/ OTHER INFORMATION: Description of Unknown Organism:Syndecan DNA
US-09-427-261-3

Query Match
Best Local Similarity 1.8%; Score 24; DB 4; Length 2193;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1304 TGTCAAAAAAAAAAAAAAAAAAAAAA 1327
Db 2159 TGTCAAAAAAAAAAAAAAAAAAAAAA 2182

RESULT 22
US-09-539-333D-222
/ Sequence 22, Application US/09539333D
/ Patent No. 6476208
/ GENERAL INFORMATION:
/ APPLICANT: Cohen, Daniel
/ APPLICANT: Blumenfeld, Marta
/ APPLICANT: Chumakov, Ilya
/ APPLICANT: Bouqueloret, Lydie
/ APPLICANT: Bihain, Bernard
/ APPLICANT: Essioux, Laurent
/ TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
/ FILE REFERENCE: GENSET.047AUS
/ CURRENT APPLICATION NUMBER: US/09/539,333D
/ CURRENT FILING DATE: 2000-03-30
/ PRIOR APPLICATION NUMBER: US 60/126,903
/ PRIOR FILING DATE: 1999-03-30
/ PRIOR APPLICATION NUMBER: US 60/131,971
/ PRIOR FILING DATE: 1999-04-30
```

```
/ PRIOR APPLICATION NUMBER: US 60/132,065
/ PRIOR FILING DATE: 1999-04-30
/ PRIOR APPLICATION NUMBER: US 60/143,928
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: US 60/145,915
/ PRIOR FILING DATE: 1999-07-27
/ PRIOR APPLICATION NUMBER: US 60/146,453
/ PRIOR FILING DATE: 1999-07-29
/ PRIOR APPLICATION NUMBER: US 60/146,452
/ PRIOR FILING DATE: 1999-07-29
/ PRIOR APPLICATION NUMBER: US 60/162,288
/ PRIOR FILING DATE: 1999-10-28
/ PRIOR APPLICATION NUMBER: US 09/416,384
/ PRIOR FILING DATE: 1999-10-12
/ NUMBER OF SEQ ID NOS: 231
/ SOFTWARE: Patent.pm
/ SEQ ID NO 222
/ LENGTH: 3001
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
/ FEATURE:
/ NAME/KEY: allele
/ LOCATION: 1501
/ OTHER INFORMATION: 99-27349-267 : polymorphic base G or A
/ FEATURE:
/ NAME/KEY: misc binding
/ LOCATION: 1502..1521
/ OTHER INFORMATION: 99-27349-267.misl, complement
/ FEATURE:
/ NAME/KEY: misc binding
/ LOCATION: 1482..1500
/ OTHER INFORMATION: 99-27349-267.mis2
/ FEATURE:
/ NAME/KEY: primer bind
/ LOCATION: 1748..1767
/ OTHER INFORMATION: upstream amplification primer, complement
/ FEATURE:
/ NAME/KEY: primer bind
/ LOCATION: 1337..1355
/ OTHER INFORMATION: downstream amplification primer
/ FEATURE:
/ NAME/KEY: misc binding
/ LOCATION: 1489..1513
/ OTHER INFORMATION: 99-27349-267 probe
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 182,848,1501,2206,2397
/ OTHER INFORMATION: n=a, g, c or t
US-09-539-333D-222

Query Match
Best Local Similarity 1.8%; Score 24; DB 4; Length 3001;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1304 TGTCAAAAAAAAAAAAAAAAAAAAAA 1327
Db 2367 TGTCAAAAAAAAAAAAAAAAAAAAAA 2390

RESULT 23
US-09-112-096-28
/ Sequence 28, Application US/09112096
/ Patent No. 6194152
/ GENERAL INFORMATION:
/ APPLICANT: Reiner Laus
/ APPLICANT: Michael H. Shapiro
/ APPLICANT: Larisa Tsavaler
/ TITLE OF INVENTION: Prostate Tumor Polynucleotide and
/ FILE REFERENCE: 7636-0015.30
/ CURRENT APPLICATION NUMBER: US/09/112,096
/ CURRENT FILING DATE: 1998-07-09
/ EARLIER APPLICATION NUMBER: 60/056,110
```

```
; EARLIER FILING DATE: 1997-08-20
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 3848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-112-096-28

Query Match      1.8%; Score 24; DB 3; Length 3848;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1304 TGTCAAAAAAAAAAAAAAAAAAAAA 1327
Db 3780 TGTCAAAAAAAAAAAAAAAAAAAAA 3803

RESULT 24
US-08-381-691-16/c
; Sequence 16, Application US/08381691
; Patent No. 5852224
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Alpha-lac Albumin Gene Constructs
; NUMBER OF SEQUENCES: 17
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/381,691
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3952 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-381-691-16

Query Match      1.8%; Score 24; DB 2; Length 3952;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1304 TGTCAAAAAAAAAAAAAAAAAAAAA 1327
Db 3856 TGTCAAAAAAAAAAAAAAAAAAAAA 3833

RESULT 25
US-09-112-096-14
; Sequence 14, Application US/09112096
; Patent No. 6194152
; GENERAL INFORMATION:
; APPLICANT: Reiner Laus
; APPLICANT: Michael H. Shapiro
; APPLICANT: Larisa Tsavaler
; TITLE OF INVENTION: Prostate Tumor Polynucleotide and
; FILE REFERENCE: 7636-0015.30
; CURRENT APPLICATION NUMBER: US/09/112,096
; CURRENT FILING DATE: 1998-07-09
; EARLIER APPLICATION NUMBER: 60/056,110
; EARLIER FILING DATE: 1997-08-20
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 5668
; TYPE: DNA

; ORGANISM: Homo sapiens
US-09-112-096-14

Query Match      1.8%; Score 24; DB 3; Length 5668;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1304 TGTCAAAAAAAAAAAAAAAAAAAAA 1327
Db 5600 TGTCAAAAAAAAAAAAAAAAAAAAA 5623

RESULT 26
US-09-636-215-777
; Sequence 777, Application US/09636215
; Patent No. 6620922
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.42717C17
; CURRENT APPLICATION NUMBER: US/09/636,215
; CURRENT FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 852
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 777
; LENGTH: 5668
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-636-215-777

Query Match      1.8%; Score 24; DB 4; Length 5668;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1304 TGTCAAAAAAAAAAAAAAAAAAAAA 1327
Db 5600 TGTCAAAAAAAAAAAAAAAAAAAAA 5623

RESULT 27
US-09-685-166A-777
; Sequence 777, Application US/09685166A
; Patent No. 6630305
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqui
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
```

```
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C21
; CURRENT APPLICATION NUMBER: US/09/685,166A
; CURRENT FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 898
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 777
; LENGTH: 5668
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-685-166A-777

Query Match      1.8%; Score 24; DB 4; Length 5668;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1304 TGTCAAAAAAAAAAAAAAAAAAAAAA 1327
DB 5600 TGTCAAAAAAAAAAAAAAAAAAAAAA 5623

RESULT 28
US-09-810-671-3/c
; Sequence 3, Application US/09810671
; Patent No. 6455291
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758
; CURRENT APPLICATION NUMBER: US/09/810,671
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 21234
; TYPE: DNA
; ORGANISM: Human
US-09-810-671-3

Query Match      1.8%; Score 24; DB 4; Length 21234;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1304 TGTCAAAAAAAAAAAAAAAAAAAAAA 1327
DB 20918 TGTCAAAAAAAAAAAAAAAAAAAAAA 20895

RESULT 29
US-10-109-854-3/c
; Sequence 3, Application US/10109854
; Patent No. 6630337
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000758DIV
; CURRENT APPLICATION NUMBER: US/10/109,854
; CURRENT FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/227,470
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 09/810,671
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 5
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 21234
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-109-854-3

Query Match      1.8%; Score 24; DB 4; Length 21234;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1304 TGTCAAAAAAAAAAAAAAAAAAAAAA 1327
DB 20918 TGTCAAAAAAAAAAAAAAAAAAAAAA 20895

RESULT 30
US-09-245-281-44/c
; Sequence 44, Application US/09245281
; Patent No. 6369196
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 07334/118001
; CURRENT APPLICATION NUMBER: US/09/245,281
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: US 09/207,359
; EARLIER FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: US 09/099,041
; EARLIER FILING DATE: 1998-06-17
; EARLIER APPLICATION NUMBER: US 09/019,942
; EARLIER FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 32042
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-281-44

Query Match      1.8%; Score 24; DB 4; Length 32042;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1304 TGTCAAAAAAAAAAAAAAAAAAAAAA 1327
DB 11653 TGTCAAAAAAAAAAAAAAAAAAAAAA 11630

RESULT 31
US-09-340-620A-63/c
; Sequence 63, Application US/09340620A
; Patent No. 6482933
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
; FILE REFERENCE: 07334-124001
; CURRENT APPLICATION NUMBER: US/09/340,620A
; CURRENT FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/245,281
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: US 09/207,359
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 09/099,041
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: US 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 32042
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-340-620A-63

Query Match      1.8%; Score 24; DB 4; Length 32042;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1304 TGTCAAAAAAAAAAAAAAAAAAAAAA 1327
Db 11653 TGTCAAAAAAAAAAAAAAAAAAAAAA 11630

RESULT 32
US-09-146-053-3/c
; Sequence 3, Application US/09146053A
; Patent No. 6399349
; GENERAL INFORMATION:
; APPLICANT: Ryan, James W.
; APPLICANT: Sprinkle, Terry Joe Curtis
; APPLICANT: Venema, Richard C.
; TITLE OF INVENTION: Human Aminopeptidase P Gene
; FILE REFERENCE: MCG103
; CURRENT APPLICATION NUMBER: US/09/146.053A
; CURRENT FILING DATE: 1998-09-02
; EARLIER APPLICATION NUMBER: 60/057,854
; EARLIER FILING DATE: 1997-09-02
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 50000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-146-053-3

Query Match      1.8%; Score 24; DB 4; Length 50000;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1304 TGTCAAAAAAAAAAAAAAAAAAAAAA 1327
Db 12947 TGTCAAAAAAAAAAAAAAAAAAAAAA 12924

RESULT 33
US-09-751-389-3/c
; Sequence 3, Application US/09751389
; Patent No. 6630334
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01067
; CURRENT APPLICATION NUMBER: US/09/751,389
; CURRENT FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 786431
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(786431)
; OTHER INFORMATION: n = A,T,C or G
US-09-751-389-3

Query Match      1.8%; Score 24; DB 4; Length 786431;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1304 TGTCAAAAAAAAAAAAAAAAAAAAAA 1327
Db 12947 TGTCAAAAAAAAAAAAAAAAAAAAAA 12924

RESULT 34
US-09-306-290-15/c
; Sequence 15, Application US/09306290
; Patent No. 6221635
; GENERAL INFORMATION:
; APPLICANT: Rovera, Giovanni
; APPLICANT: Mukhopadhyay, Sunil
; TITLE OF INVENTION: METHODS FOR SOLID-PHASE AMPLIFICATION OF DNA TEMPLATE
; TITLE OF INVENTION: (SPADT) USING MULTIBARRAYS
; FILE REFERENCE: 09924-10
; CURRENT APPLICATION NUMBER: US/09/306,290
; CURRENT FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 40
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer RGP
; OTHER INFORMATION: RR915
US-09-306-290-15

Query Match      1.7%; Score 23; DB 3; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1305 GTCACAAAAAAAAAAAAAAAAAAAAA 1327
Db 23 GTCACAAAAAAAAAAAAAAAAAAAAA 1

RESULT 35
US-09-621-976-10956
; Sequence 10956, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 10956
; LENGTH: 67
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-10956

Query Match      1.7%; Score 23; DB 4; Length 67;
Best Local Similarity 100.0%; Pred. No. 0.45;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1305 GTCACAAAAAAAAAAAAAAAAAAAAA 1327
Db 32 GTCACAAAAAAAAAAAAAAAAAAAAA 54

RESULT 36
US-09-621-976-9086
; Sequence 9086, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
```



```
/ FILE REFERENCE: GENSET.054PR2
/ CURRENT APPLICATION NUMBER: US/09/621.976
/ CURRENT FILING DATE: 2000-07-21
/ NUMBER OF SEQ ID NOS: 19335
/ SOFTWARE: Patent.pm
/ SEQ ID NO 9086
/ LENGTH: 97
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-621-976-9086

Query Match      1.7%; Score 23; DB 4; Length 97;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1305 GTCAAAAAAAAAAAAAAAAAAAAA 1327
Db 35 GTCAAAAAAAAAAAAAAAAAAAAA 57

RESULT 37
US-09-091-725-28
/ Sequence 28, Application US/09091725
/ Patent No. 6329141
/ GENERAL INFORMATION:
/ APPLICANT:
/ TITLE OF INVENTION: Improved methods for transforming Phaffia
/ TITLE OF INVENTION: and recombinant DNA for use therein
/ NUMBER OF SEQUENCES: 51
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Morrison & Foerster llp
/ STREET: 2000 Pennsylvania Avenue, N.W.
/ CITY: Washington
/ STATE: DC
/ COUNTRY: United States of America
/ ZIP: 20006-1888
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25 (BFO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/091,725
/ FILING DATE: 23-DEC-1996
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: EP 95203620.0
/ FILING DATE: 22-DEC-1995
/ APPLICATION NUMBER: EP 96200943.7
/ FILING DATE: 11-APR-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: E. Victor Donahue
/ REGISTRATION NUMBER: 35,492
/ INFORMATION FOR SEQ ID NO: 28:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 307 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:
/ ORGANISM: Phaffia rhodozyma
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 3..227
/ OTHER INFORMATION: /product= "PRCDNA18"
US-09-091-725-28

Query Match      1.7%; Score 23; DB 4; Length 307;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1305 GTCAAAAAAAAAAAAAAAAAAAAA 1327
Db 261 GTCAAAAAAAAAAAAAAAAAAAAA 283
```

```
RESULT 38
US-09-621-976-15342
/ Sequence 15342, Application US/09621976
/ Patent No. 6639063
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, J.B.
/ APPLICANT: Jobert, S. Y.
/ APPLICANT: Giordano, J.Y.
/ TITLE OF INVENTION: ESTs and Encoded Human Proteins.
/ FILE REFERENCE: GENSET.054PR2
/ CURRENT APPLICATION NUMBER: US/09/621,976
/ CURRENT FILING DATE: 2000-07-21
/ NUMBER OF SEQ ID NOS: 19335
/ SOFTWARE: Patent.pm
/ SEQ ID NO 15342
/ LENGTH: 350
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-621-976-15342
```

```
Query Match      1.7%; Score 23; DB 4; Length 350;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1305 GTCAAAAAAAAAAAAAAAAAAAAA 1327
Db 287 GTCAAAAAAAAAAAAAAAAAAAAA 309
```

```
RESULT 39
US-07-987-272A-13
/ Sequence 13, Application US/07987272A
/ Patent No. 5731166
/ GENERAL INFORMATION:
/ APPLICANT: Geczy, C., Simpson, R. J. and Lackmann, M
/ TITLE OF INVENTION: No. 5731166el Chemotactic Factor
/ NUMBER OF SEQUENCES: 23
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Cushman Darby & Cushman
/ STREET: 1100 New York Avenue, N. W., Ninth Floor, East Tower
/ CITY: Washington
/ STATE: D. C.
/ COUNTRY: USA
/ ZIP: 20005-3918
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/987,272A
/ FILING DATE: 05-MAR-1993
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: AU PK 2127
/ FILING DATE: 05-FEB-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: AU PK 4463
/ FILING DATE: 05-SEP-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Brinkman, David W
/ REGISTRATION NUMBER: 20,817
/ REFERENCE/DOCKET NUMBER: DMB/1925/200259
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-861 3000
/ TELEFAX: 202-822 0944
/ TELEX: 6714627 CUSH
```

NAME: Steffe, Eric K.

```
;
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0420001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 989 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 88..534
; NAME/KEY: sig_peptide
; LOCATION: 88..147
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 148..534
; US-08-874-460-1
;
; Query Match 1.7%; Score 23; DB 2; Length 989;
; Best Local Similarity 100.0%; Pred. No. 0.44;
; Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1305 GTCACAAAAA 1327
; Db 958 GTCACAAAAA 980
;
; RESULT 43
; US-09-272-162-1
; Sequence 1, Application US/09272162
; Patent No. 6503735
; GENERAL INFORMATION:
; APPLICANT: WEI, YING-FEI
; KREIDER, BRENT
; ROSEN, CRAIG
; TITLE OF INVENTION: CHEMOKINE BETA 15
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSER: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: US
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/272,162
; FILING DATE: 19-Mar-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/874,460
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Steffe, Eric K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.0420001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 989 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 88..534
; NAME/KEY: sig_peptide
; LOCATION: 88..147
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 148..534
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
; US-09-272-162-1
;
; Query Match 1.7%; Score 23; DB 4; Length 989;
; Best Local Similarity 100.0%; Pred. No. 0.44;
; Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1305 GTCACAAAAA 1327
; Db 958 GTCACAAAAA 980
;
; RESULT 44
; US-09-647-224A-5
; Sequence 5, Application US/09647224A
; Patent No. 6482631
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Gutteridge, Steven
; APPLICANT: Hitz, William D.
; APPLICANT: Maxwell, Carl A.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Tao, Yong
; APPLICANT: Vollmer, Steven J.
; TITLE OF INVENTION: TRYPTOPHAN BIOSYNTHETIC ENZYMES
; FILE REFERENCE: BB-1150-A
; CURRENT APPLICATION NUMBER: US/09/647,224A
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 60/079,386
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: PCT/US99/06046
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 1005
; TYPE: DNA
; ORGANISM: Triticum aestivum
; US-09-647-224A-5
;
; Query Match 1.7%; Score 23; DB 4; Length 1005;
; Best Local Similarity 100.0%; Pred. No. 0.44;
; Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 1303 ATGTCACAAAAA 1325
; Db 963 ATGTCACAAAAA 1005
;
; RESULT 45
; US-09-396-149-17
; Sequence 17, Application US/09396149
; Patent No. 6538176
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Replication Protein A and Use
; FILE REFERENCE: 5718-59
; CURRENT APPLICATION NUMBER: US/09/396,149
; CURRENT FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 17
; LENGTH: 1087
; TYPE: DNA
; ORGANISM: Zea mays
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Maize RPA Middle Subunit Homologue-5
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (91)...(1044)
US-09-396-149-17

Query Match      1.7%; Score 23; DB 4; Length 1087;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1304 TGTCACAAAAA 1326
Db 1065 TGTCACAAAAA 1087

RESULT 46
US-08-036-555B-135
; Sequence 135, Application US/08036555B
; Patent No. 5530109
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew; Stroobant, Paul;
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
; APPLICANT: Chen, Maio Su; Hiles, Ian
; TITLE OF INVENTION: Glial Mitogenic Factors, Their
; TITLE OF INVENTION: Preparation and Use
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/036,555B
; FILING DATE: 24-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; APPLICATION DATA:
; FILING DATE: 30-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/963,703
; FILING DATE: 03-APRIL-1992
; APPLICATION DATA:
; FILING DATE: 10-APRIL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: LUD 5250.4
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 135:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1087
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-469-569-135

; LENGTH: 1108
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-036-555B-135

Query Match      1.7%; Score 23; DB 1; Length 1108;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1305 GTCAAAAAA 1327
Db 1059 GTCAAAAAA 1081

RESULT 47
US-08-469-569-135
; Sequence 135, Application US/08469569
; Patent No. 5606032
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew; Stroobant, Paul;
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
; APPLICANT: Chen, Maio Su; Hiles, Ian
; TITLE OF INVENTION: Glial Mitogenic Factors, Their
; TITLE OF INVENTION: Preparation and Use
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,569
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 24-MAR-1993
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; APPLICATION DATA:
; FILING DATE: 30-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/963,703
; FILING DATE: 03-APRIL-1992
; APPLICATION DATA:
; FILING DATE: 10-APRIL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: LUD 5250.4
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 135:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1108
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-469-569-135
```

Query Match 1.7%; Score 23; DB 1; Length 1108;
Best Local Similarity 100.0%; Pred. No. 0.44; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0;

QY 1305 GTCACAAAAA 1327
Db 1059 GTCACAAAAA 1081

RESULT 48
US-08-249-322A-135
; Sequence 135, Application US/08249322A
; Patent No. 5716930
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew; Stroobant, Paul;
; APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
; APPLICANT: Chen, Maio Su; Hiles, Ian
; TITLE OF INVENTION: Glial Mitogenic Factors, Their
; PREPARATION AND USE
; NUMBER OF SEQUENCES: 184
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felle & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/249,322A
; FILING DATE: 26-MAY-1994
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 24-MAR-1993
; APPLICATION DATA:
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; APPLICATION DATA:
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; APPLICATION DATA:
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APRIL-1992

; APPLICATION NUMBER: U.K. 91 07566.3
; FILING DATE: 10-APRIL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: LUD 250.4
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 135:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1108
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-249-322A-135

Query Match 1.7%; Score 23; DB 1; Length 1108;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1305 GTCACAAAAA 1327
Db 1059 GTCACAAAAA 1081

RESULT 49
US-08-469-526A-135
; Sequence 135, Application US/08469526A
; Patent No. 5792849
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew
; APPLICANT: Stroobant, Paul
; APPLICANT: Minghetti, Luisa
; APPLICANT: Waterfield, Michael
; APPLICANT: Marchionni, Mark
; APPLICANT: Chen, Maio Su
; APPLICANT: Hiles, Ian
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
; PREPARATION AND USE
; NUMBER OF SEQUENCES: 187
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,526A
; FILING DATE: 06 June 1995
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 24-MAR-1993
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 03-JUN-1992
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APRIL-1992
; APPLICATION NUMBER: U.K. 91 07566.3
; FILING DATE: 10-APR-1991

; ATTORNEY/AGENT INFORMATION:
; NAME: Bleker-Brady, Kristina
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 04585/00200A
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; INFORMATION FOR SEQ ID NO: 135:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1108 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 8..778
; OTHER INFORMATION:
US-08-469-526A-135

Query Match 1.7%; Score 23; DB 1; Length 1108;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1305 GTCACAAAAA 1327
Db 1059 GTCACAAAAA 1081

RESULT 50

US-08-734-591A-135
; Sequence 135 Application US/08734591A
; Patent No. 5854220
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew
; APPLICANT: Stroobant, Paul
; APPLICANT: Minghetti, Luisa
; APPLICANT: Waterfield, Michael
; APPLICANT: Hiles, Ian
; APPLICANT: Marchionni, Mark
; APPLICANT: Chen, Mario
; TITLE OF INVENTION: GLIAL MITOGENIC FACTORS, THEIR
; NUMBER OF SEQUENCES: 187
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM Compatible Pentium
; OPERATING SYSTEM: Windows95
; SOFTWARE: WordPerfect (Version 7.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/734,591A
; FILING DATE: 22-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/470,335
; FILING DATE: 06-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/036,555
; FILING DATE: 03-MAR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/965,173
; FILING DATE: 23-OCT-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/940,389
; FILING DATE: 03-SEP-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/907,138
; FILING DATE: 30-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/863,703
; FILING DATE: 03-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: UK 91 075666.3
; FILING DATE: 10-APR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Bieker-Brady, Kristina
; REGISTRATION NUMBER: 39,109
; REFERENCE/DOCKET NUMBER: 04585/00200P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 428-0200
; TELEFAX: (617) 428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 135:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1108 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 8...778
; OTHER INFORMATION:
US-08-734-591A-135

Query Match 1.7%; Score 23; DB 2; Length 1108;
Best Local Similarity 100.0%; Pred. No. 0.44; Indels 0; Gaps 0;
Matches 23; Conservative 0; Mismatches 0;
QY 1305 GTCACAAAAA 1327
Db 1059 GTCACAAAAA 1081

Search completed: September 17, 2004, 05:42:46
Job time : 129 secs